



**Charakterisierung eines minimierten eukaryoten
Cytoplasmas - das periplastidäre Kompartiment der
Diatomee *Phaeodactylum tricornutum***

Dissertation

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem Fachbereich Biologie

der Philipps-Universität Marburg

vorgelegt von

M.Sc.

Daniel Moog

aus Marburg

Marburg/Lahn 2012

Vom Fachbereich Biologie der Philipps-Universität Marburg (Hochschulkennziffer 1180)
als Dissertation angenommen am: 19.12.2012

Erstgutachter: Prof. Dr. Uwe-G. Maier

Zweitgutachter Prof. Dr. Andreas Brune

Prof. Dr. Renate Renkawitz-Pohl

Prof. Dr. Alfred Batschauer

Tag der Disputation am: 21.12.2012

Publikationen

Teile dieser Arbeit wurden publiziert in:

Moog D, Stork S, Zauner S, Maier UG (2011) *In silico* and *in vivo* investigations of proteins of a minimized eukaryotic cytoplasm. *Genome Biol Evol.* 3: 375-382.

Stork S* and **Moog D***, Przyborski JM, Wilhelmi I, Zauner S, Maier UG (2012) Distribution of the SELMA translocon in secondary plastids of red algal origin and predicted uncoupling of ubiquitin-dependent translocation from degradation. *Eukaryot Cell*, doi:10.1128/EC.00183-12.

*Gleicher Beitrag zum Manuskript

Weitere Publikationen:

Felsner G, Sommer MS, Gruenheit N, Hempel F, **Moog D**, Zauner S, Martin W, Maier UG (2010) ERAD components in organisms with complex red plastids suggest recruitment of a preexisting protein transport pathway for the periplastid membrane. *Genome Biol Evol.* 3: 140-150.

Mernberger M, **Moog D**, Stork S, Zauner S, Maier UG, Hüllermeier E (2011) Prediction of protein localization for specialized compartments using time series kernels. *Proc. GCB, LNI.*

Inhaltsverzeichnis

Inhaltsverzeichnis.....	I
Abbildungsverzeichnis	V
Tabellenverzeichnis.....	VI
Abkürzungsverzeichnis	VII
1. Einleitung	1
1.1 Evolution primärer und komplexer Plastiden.....	1
1.1.1 Primäre Endosymbiose - die Entstehung photosynthetischer Eukaryoten	1
1.2.2 Eukaryote Endosymbiose - die Entstehung komplexer Plastiden	3
1.2. Proteinimport in primäre und komplexe Plastiden.....	7
1.2.1 Proteinimport in primäre Plastiden.....	8
1.2.2 Proteinimport in komplexe Plastiden rhodophytischen Ursprungs	9
1.3 Das periplastidäre Kompartiment – ein minimiertes eukaryotes Cytoplasma.....	14
1.4. Triosephosphat/Phosphat-Translokatoren in primären und komplexen Plastiden	19
1.5 <i>Phaeodactylum tricornutum</i> als Modellorganismus.....	22
2. Zielsetzung	24
3. Ergebnisse	25
3.1. Charakterisierung des periplastidären Kompartiments von <i>Phaeodactylum tricornutum</i>	25
3.1.1 Proteinimport - SELMA-Faktoren	28
3.1.1.1 Interaktionsstudien mit sCdc48-1/2 und sPub mittels Split-GFP System	30
3.1.2 Molekulare Chaperone	31
3.1.3 Proteasomale Komponenten und Proteasen	34
3.1.4 Plastidenteilungsfaktoren	39
3.1.4.1 Bioinformatische Analysen und <i>in vivo</i> Lokalisationen putativer Komponenten des Teilungsapparats der komplexen Plastide	39
3.1.4.2 Analyse des Glycosylierungsstatus von sPdr1	43
3.1.5 Lipidtransfer, Fettsäurebiosynthese und Glycero(phospho)lipidmetabolismus	45
3.1.6 Metabolismus - Carboanhydrasen	46
3.1.7 Homologe zu Nucleomorph-kodierten Proteinen der Cryptophyte <i>Guillardia theta</i>	48
3.1.8 Weitere Faktoren	50
3.2 Homologe zu PPC-Proteinen aus <i>P. tricornutum</i> in anderen Organismen mit komplexer Plastide rhodophytischen Ursprungs und Rotalgen	53
3.3 Putative Triosephosphat/Phosphat-Translokatoren in <i>P. tricornutum</i>	56
3.3.1 Bioinformatische Analysen und <i>in vivo</i> Lokalisation putativer TPTs in <i>P. tricornutum</i>	57

3.3.2 Untersuchung der Zielmembranen und Orientierungen putativer TPTs in der komplexen Plastide von <i>P. tricornutum</i> über das Self-Assembling-GFP System	61
3.3.3 Analyse der Membranintegration putativer TPTs in den Hüllmembranen der komplexen Plastide von <i>P. tricornutum</i>	64
4. Diskussion	66
4.1 Neue Proteine und Funktionen, sowie abwesende Prozesse im PPC von <i>P. tricornutum</i>	66
4.1.1 SELMA und faltungsrelevante Faktoren	67
4.1.2 Besitzt das PPC eine proteolytische Aktivität?	71
4.1.3 Das PPC beherbergt keine typischen Vesikel- und Cytoskelettelemente	75
4.1.4 Komplexe Plastide - simple Teilung?	79
4.1.5 Metabolische Prozesse im PPC sind begrenzt	83
4.1.6 Homöostase und Regulationsmechanismen im PPC	85
4.2 Das PPC der Diatomee <i>Phaeodactylum tricornutum</i>	87
4.3 Das PPC von <i>P. tricornutum</i> im evolutionären Kontext	89
4.4 Ist eine <i>in silico</i> Charakterisierung des PPCs möglich?	92
4.5 Triosephosphat/Phosphat Translokatoren als metabolische Verbindung zwischen Wirt und Symbiont?	94
5. Zusammenfassung	98
5.1 Abstract	99
6. Material und Methoden	100
6.1 Material	100
6.1.1 Chemikalien	100
6.1.2 Enzyme	100
6.1.3 Molekularbiologische Kits	100
6.1.4 Geräte	101
6.1.5 Vektoren	102
6.1.6 Antikörper	102
6.1.7 Organismen	102
6.1.8 Oligonukleotide (Primer)	102
6.1.9 DNA- und Proteinmolekulargewichtsstandards	106
6.1.10 Hardware, Software und Internetanwendungen	107
6.1.11 Puffer und Lösungen	108
6.2 Methoden	108
6.2.1 Kultur- und Stammhaltung	108
6.2.1.1 Kultur- und Stammhaltung der Diatomee <i>Phaeodactylum tricornutum</i>	108
6.2.1.2 Kultur- und Stammhaltung von <i>Escherichia coli</i> TOP10	109
6.2.2 Nukleinsäureanalytik	110

6.2.2.1 Isolation von Gesamt-DNA (gDNA) und -RNA (gRNA) aus <i>Phaeodactylum tricornutum</i>	110
6.2.2.2 DNaseI-Behandlung isolierter gRNA.....	110
6.2.2.3 Reverse Transkription (RT-PCR).....	110
6.2.2.4 Amplifikation von Nukleinsäuren - Polymerase-Kettenreaktion	110
6.2.2.5 Agarose-Gelelektrophorese und Elution von DNA Fragmenten.....	111
6.2.2.6 DNA-Restriktion	112
6.2.2.7 Ligation von DNA-Fragmenten	112
6.2.2.8 Transformation von RbCl ₂ -kompetenten <i>Escherichia coli</i> TOP10-Zellen.....	112
6.2.2.9 Plasmidisolation aus <i>Escherichia coli</i> mittels alkalischer Lyse	113
6.2.2.10 DNA-Sequenzierung (ABI-377)	114
6.2.2.11 Biolistische Transformation der Diatomee <i>Phaeodactylum tricornutum</i>	115
6.2.3 Proteinanalytik.....	116
6.2.3.1 Herstellung von Proteingesamtextrakt aus <i>Phaeodactylum tricornutum</i>	116
6.2.3.2 Protein-Konzentrationsbestimmung mittels Amidoschwarz 10B	117
6.2.3.3 SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE).....	118
6.2.3.4 Coomassie-Färbung von SDS-Gelen.....	118
6.2.3.5 Western Blot und Immunodetektion von Proteinen	118
6.2.3.6 Immunopräzipitation von rekombinanten Proteinen aus <i>Phaeodactylum tricornutum</i>	120
6.2.3.7 Carbonatextraktion	120
6.2.3.8 Färbung zum Nachweis von Glycoproteinen (<i>Glyco-Stain</i>).....	121
6.2.3.9 PNGase F-Behandlung von Glycoproteinen	121
6.2.3.10 <i>In vivo</i> Lokalisationsstudien - Expression rekombinanter GFP-Reporterfusionsproteine in <i>Phaeodactylum tricornutum</i>	121
6.2.3.10.1 Präsequenz- und „volle Länge“-GFP Lokalisationsstudien	121
6.2.3.10.2 Self-Assembling GFP - <i>in vivo</i> Analysen zur Topologie von Membranproteinen	122
6.2.3.10.3 Split-GFP - Analysen von <i>in vivo</i> Proteininteraktionen	123
6.2.4 Konfokale Laser Scanning Mikroskopie	124
6.2.5 Bioinformatische (<i>in silico</i>) Analysen.....	125
7. Literatur	127
8. Anhang	141
8.1. Tabellarische Zusammenfassung der Ergebnisse der kombinierten <i>in silico/in vivo</i> Ansätze zur Identifizierung PPC-lokalisierter Proteine in <i>P. tricornutum</i>	141
8.2. BTS-Sequenzen PPC-lokalisierter Faktoren aus <i>P. tricornutum</i>	160
8.3. Alignments	164
8.3.1. sCdc28-1/2 und hCdc48	164

8.3.2 sDrp und Cm_Dnm2(DRP5B)	165
8.3.3 sPdr1 und Cm_PDR1	165
8.3.4 sORF139 und der Nucleomorph-kodierte ORF139 aus <i>G. theta</i> mit dem Homolog aus <i>C. merolae</i>	165
8.3.5 sORF261 und Homologe aus <i>G. theta</i> und <i>C. merolae</i>	166
8.3.6 sORF532a und Homologe aus <i>G. theta</i> und <i>C. merolae</i>	166
8.3.7 sORF534 und Homologe aus <i>G. theta</i> und <i>C. merolae</i>	166
8.4 Proteinexpressionstests	167
8.5 <i>In silico</i> Analysen zu putativ PPC-lokalisierten Proteinen in Organismen mit komplexer Plastide rhodophytischen Ursprungs und Rotalgen	168
Danksagung	173
Curriculum Vitae	174
Erklärung	175

Abbildungsverzeichnis

Abb. 1.1 Primäre Endosymbiose	2
Abb. 1.2 Die Entstehung von Organismen mit komplexen Plastiden rhodophytischen Ursprungs	4
Abb. 1.3 Import nucleuskodierter Proteine in komplexe Plastiden rhodophytischen Ursprungs am Beispiel der Diatomee <i>P. tricornutum</i>	12
Abb. 1.4 Morphologie des periplastidären Kompartiments (PPC) von <i>P. tricornutum</i>	17
Abb. 1.5 Das PPC der Diatomee <i>Phaeodactylum tricornutum</i>	18
Abb. 1.6 Importstudien von GFP-Fusionsproteinen in der komplexen Plastide von <i>P. tricornutum</i>	23
Abb. 3.1 Flussdiagramm zu den beiden kombinierten <i>in silico/in vivo</i> Ansätzen zur Identifizierung periplastidärer Proteine in <i>P. tricornutum</i>	26
Abb. 3.2 Domänenalignments der Proteinprimärsequenzen von wirts- und symbiontspezifischen Cdc48- und Pub-Proteinen aus <i>P. tricornutum</i>	28
Abb. 3.3 <i>In vivo</i> Lokalisation von sCdc48-2 und sPub im PPC von <i>P. tricornutum</i>	29
Abb. 3.4 <i>In vivo</i> Interaktionsstudien von sCdc48-1/2 und sPub im PPC von <i>P. tricornutum</i> über das Split-GFP System	31
Abb. 3.5 Detektierte Domänen putativer PPC-Chaperone aus <i>P. tricornutum</i>	32
Abb. 3.6 <i>In vivo</i> Lokalisation neuer Chaperone im PPC von <i>P. tricornutum</i>	33
Abb. 3.7 <i>In vivo</i> Lokalisation proteasomaler 20S-Faktoren im PPC von <i>P. tricornutum</i>	35
Abb. 3.8 Domänenkomposition der Tlp-Proteine aus <i>P. tricornutum</i>	37
Abb. 3.9 Primärsequenz-Alignment der aktiven Zentren und stark konservierten Sequenzbereiche der Tlp-Proteine aus <i>P. tricornutum</i>	37
Abb. 3.10 <i>In vivo</i> Lokalisation symbiontischer Trypsin-ähnlicher Serin-Proteasen (sTlp) im PPC von <i>P. tricornutum</i>	38
Abb. 3.11 Domänenalignment putativer Komponenten des Teilungsapparates der komplexen Plastide von <i>P. tricornutum</i>	40
Abb. 3.12 <i>In vivo</i> Lokalisation putativer Komponenten des Teilungsapparates der komplexen Plastide im PPC von <i>P. tricornutum</i>	42
Abb. 3.13 Analyse des Glycosylierungsstatus des PPC-lokalisierten sPdr1_FL-eGFP aus <i>P. tricornutum</i>	44
Abb. 3.14 <i>In vivo</i> Lokalisation von sSec14 im PPC von <i>P. tricornutum</i>	46
Abb. 3.15 <i>In vivo</i> Lokalisation putativer α -Carboanhydrasen im PPC von <i>P. tricornutum</i>	47
Abb. 3.16 <i>In vivo</i> Lokalisation homologer Proteine zu Nucleomorph-kodierten ORFs der Cryptophyte <i>Guillardia theta</i> im PPC von <i>P. tricornutum</i>	50
Abb. 3.17 <i>In vivo</i> Lokalisation von sPrp, sP4h, sPel und sSmc im PPC von <i>P. tricornutum</i>	51

Abb. 3.18 Identifizierte Proteindomänen der PPC-lokalisierten Proteine sP4h, sPel und sSmc aus <i>P. tricornutum</i>	52
Abb. 3.19 Domänen-Alignment putativer Triosephosphat/Phosphat-Translokatoren aus <i>P. tricornutum</i>	57
Abb. 3.20 <i>In vivo</i> Lokalisation putativer TPTs in der komplexen Plastide von <i>P. tricornutum</i>	60
Abb. 3.21 <i>In vivo</i> Lokalisation weiterer TPTs außerhalb der komplexen Plastide von <i>P. tricornutum</i>	61
Abb. 3.22 Analyse der Lokalisation und Membranorientierung putativer TPTs in der komplexen Plastide von <i>P. tricornutum</i> mit dem Self-Assembling-GFP System	64
Abb. 3.23 Western Blot nach Carbonatextraktion mit den TPTs der komplexen Plastide von <i>P. tricornutum</i>	65
Abb. 4.1 Primärsequenz-Alignment des C-Terminus von Cdc48-Proteinen aus <i>P. tricornutum</i>	69
Abb. 4.2 Das PPC der Diatomee <i>Phaeodactylum tricornutum</i>	88
Abb. 4.3 Modell zur Lokalisation putativer Triosephosphat/Phosphat-Translokatoren in der komplexen Plastide von <i>P. tricornutum</i>	97
Abb. 8.1 TargetP-Analyse der putativen TPL-Sequenzen <i>in vivo</i> lokalisierter PPC-Proteine aus <i>P. tricornutum</i>	162
Abb. 8.2 ChloroP-Analyse der putativen TPL-Sequenzen <i>in vivo</i> lokalisierter PPC-Proteine aus <i>P. tricornutum</i>	163
Abb. 8.3 Alignment mit relativen Aminosäure-Häufigkeiten der TPL-Sequenzen aller verifizierter PPC-Proteine in <i>P. tricornutum</i>	164
Abb. 8.4 Expressionstest von sPdr1_FL-eGFP in <i>P. tricornutum</i> unter Nitrat-induzierbarem Promotor	167
Abb. 8.5 Expressionstest von TPT1-4_FL-eGFP in <i>P. tricornutum</i> über pPha-T1	167

Tabellenverzeichnis

Tab. 3.1 Vergleich der PPC-Proteine aus <i>P. tricornutum</i> mit putativen Homologen aus anderen Organismen mit komplexer Plastide rhodophytischen Ursprungs und rezenten Rotalgen	55
Tab. 3.2 Detektierte Präsequenzen putativer TPTs aus <i>P. tricornutum</i>	58
Tab. 8.1 Ergebnisse der kombinierten <i>in silico/in vivo</i> Ansätze zur Identifizierung PPC-lokalisierter Proteine in <i>P. tricornutum</i>	141
Tab. 8.2 BTS-Sequenzen PPC-lokalisierter Faktoren aus <i>P. tricornutum</i> ermittelt über Signalpeptidvorhersage (SignalP) und Detektion N-terminaler Extensionen (BlastP)	160
Tab. 8.3 <i>In silico</i> Analysen zu putativ PPC-lokalisierten Proteinen in Organismen mit komplexer Plastide rhodophytischen Ursprungs und Rotalgen	168

Abkürzungsverzeichnis

°C	Grad Celsius	OEM	<i>plastid outer envelope membrane</i>
AS	Aminosäure(n)	ORF	<i>open reading frame</i>
ATP	Adenosintriphosphat	PPC	<i>periplastidal compartment</i> (periplastidäres Kompartiment)
bp	<i>base pair(s)</i>	PPM	<i>periplastidal membrane</i> (periplastidäre Membran)
BLAST	<i>Basic local alignment search tool</i>	rER	raues ER
BTS	<i>bipartite topogenic signal</i>	RNA	Ribonukleinsäure
ca.	circa	RT	Raumtemperatur
cDNA	<i>complementary DNA</i>	s	<i>symbiont, symbiont specific</i> (symbiontspezifisch)
CDS	<i>conserved domain search</i>	sek.	Sekunde(n)
cER-(M)	<i>chloroplast endoplasmic reticulum (membrane)</i>	SELMA	<i>symbiont specific ERAD-like machinery</i>
d.h.	das heißt	SP	Signalpeptid
ddH ₂ O	doppelt destilliertes Wasser	TIC	<i>translocon of the inner chloroplast envelope membrane</i>
dH ₂ O	destilliertes Wasser	TMD	Transmembrandomäne
DNA	Desoxyribonukleinsäure	TPT	Triosephosphat/Phosphat-translokator
eGFP	<i>enhanced green fluorescent protein</i>	TOC	<i>translocon of the outer chloroplast envelope membrane</i>
ER	Endoplasmatisches Reticulum	TP	Transitpeptid
ERAD	ER-assoziierte Degradation	TPL	<i>transit peptide-like sequence</i> (Transitpeptid-ähnliche Sequenz)
EST	<i>expressed sequence tag</i>	U	Unit
et al.	et alii (und andere)	u.a.	unter anderem
FL	<i>full length</i> (volle Länge)	u.U.	unter Umständen
g	Gramm	ü.N.	über Nacht
GFP	<i>green fluorescent protein</i> (grün fluoreszierendes Protein)	V	Volt
ggf.	gegebenenfalls	vgl.	Vergleich
h	<i>host, host specific</i> (wirtsspezifisch)	Vol.	Volumen
IEM	<i>plastid inner envelope membrane</i>	v/v	<i>volume per volume</i>
IMS	<i>intermembrane space</i>	w/v	<i>weight per volume</i>
KLSM	konfokales Laser Scanning Mikroskop	z.B.	zum Beispiel
l	Liter	z.T.	zum Teil
M	Molar (mol/l)	µl	Mikroliter (10 ⁻⁶ l)
mA	Milliampere (10 ⁻³ A)	µg	Mikrogramm (10 ⁻⁶ g)
mg	Milligramm (10 ⁻⁶ kg)	µM	Mikromol (10 ⁻⁶ M)
min	Minute(n)		
ml	Milliliter (10 ⁻³ l)		
mM	Millimolar (10 ⁻³ M)		
NADPH	Nicotinamidadenindinucleotid-phosphat		
NTP	Nukleosidtriphosphat		

1. Einleitung

1.1 Evolution primärer und komplexer Plastiden

Bereits früh in der Erdgeschichte machten sich Organismen die Energie der Sonne für metabolische Zwecke zu Nutze. Es wird vermutet, dass sich der Prozess oxydativer Photosynthese, die biochemische Transformation von Wasser und Kohlendioxid zu Sauerstoff und energiereichen organischen Verbindungen durch Lichtenergie, bereits vor ca. 3,5 Milliarden Jahren in Vorfahren der heutigen Cyanobakterien etablierte (Blankenship, 1992). Die kontinuierliche Anreicherung der Atmosphäre des Planeten mit Sauerstoff stellte die Grundvoraussetzung für die Evolution komplexen Lebens dar. Die Fähigkeit Photosynthese zu betreiben blieb jedoch nicht auf prokaryote Organismen beschränkt, sondern fand auch Einzug in das Reich der Eukaryoten. Letztere entwickelten diese Kapazität allerdings nicht eigenständig, sondern durch Domestizierung eines photosynthetischen Prokaryoten, was zu den heute als Plastiden bekannten typischen Organellen photosynthetischer Eukaryoten (Algen und Pflanzen), führte. Bereits 1905 wurde von Constantin Mereschkowsky die Ähnlichkeit von Plastiden und Cyanobakterien festgestellt (Mereschkowsky, 1905), die moderne Endosymbiontentheorie wurde jedoch erst 1970 von Lynn Margulis formuliert (Margulis, 1970).

1.1.1 Primäre Endosymbiose - die Entstehung photosynthetischer Eukaryoten

Die Entstehung photosynthetischer Eukaryoten geht auf das Ereignis primärer Endosymbiose zurück bei dem ein anzestrales, freilebendes, zur oxygenen Photosynthese befähigtes Cyanobakterium von einem heterotrophen eukaryoten Wirt internalisiert und als semiautonomes Organell (primäre Plastide/Chloroplast) etabliert wurde (Cavalier-Smith, 2000; Douglas & Turner, 1991). Dieses Ereignis wird anhand „molekularer Uhren“ auf das späte Paläoproterozoikum vor ca. 1,6 Milliarden Jahren datiert (Yoon et al, 2004) und bezeichnet den Aufstieg oxygener Photosynthese in das Reich der Eukaryoten. Dabei entstand höchstwahrscheinlich zunächst eine Art zelluläre Chimäre bei der das Cytosol der Wirtszelle von den beiden Hüllmembranen des Cyanobakteriums durch die Membran einer phagotrophen Vakuole des Wirtes separiert wurde (Maier et al, 2000). Die stabile Integration

des Endosymbionten als Organell ging einher mit Reduktionsmechanismen, bei denen sowohl eine der drei umgebenden Hüllmembranen (wahrscheinlich die der phagotropen Vakuole) als auch redundantes genetisches Material des Symbionten verloren gingen (Stoebe & Maier, 2002). Außerdem fand eine massive genetische Reorganisation statt, bei der ein Großteil essentieller Gene des Endosymbionten in das Genom der Wirtszelle transferiert wurde, wodurch der Wirt nahezu vollständige genetische Kontrolle über den Endosymbionten erlangte (Maier et al, 2000; Martin et al, 2002). Die Entwicklung spezifischer Zielsteuerungssequenzen, sowie einer Translokationsmaschinerie für fortan nucleuskodierte Symbiontenproteine gilt als kritischer Schritt bei der Etablierung des Endosymbionten als Organell. Dies hat sich höchstwahrscheinlich bereits früh nach dem Ereignis der primären Endosymbiose ereignet, bevor es zu Diversifikation in die rezenten Linien der Chlorophyta (Grünalgen, Vorläufer der Landpflanzen), Rhodophyta (Rotalgen) und Glaucophyta kam (Abb 1.1) (Gould et al, 2008).

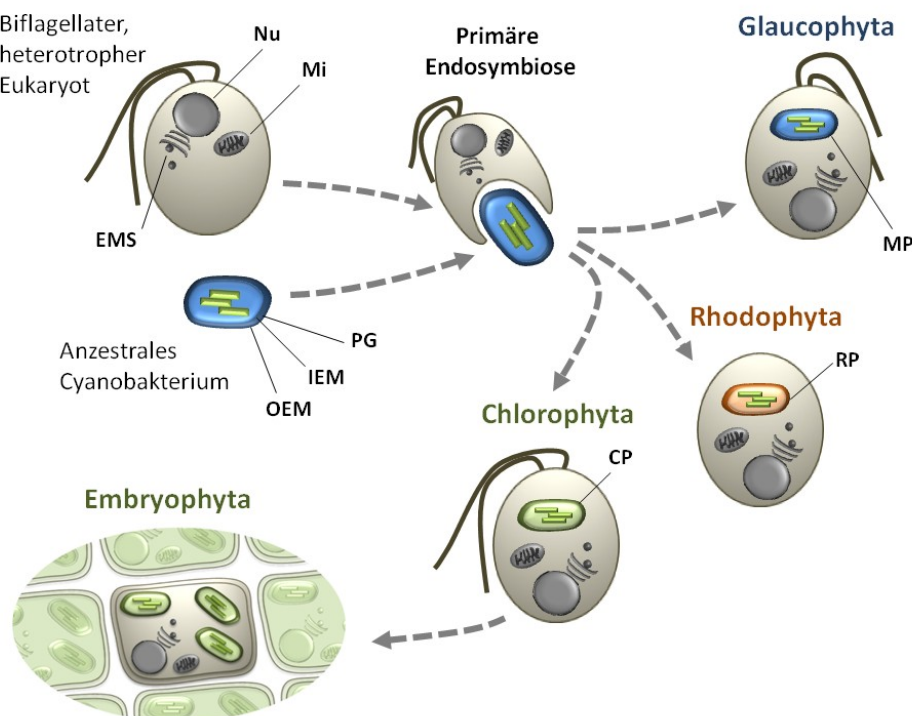


Abb. 1.1 Primäre Endosymbiose. Schematische Darstellung des monophyletischen Ereignisses der primären Endosymbiose bei dem ein Cyanobakterien-ähnlicher Vorläufer von einem heterotrophen Eukaryot internalisiert und als primäre Plastide etabliert wurde. Aus diesem Ereignis gingen die Linien der Glaucophyta, Rhodophyta und Chlorophyta, aus denen sich die Landpflanzen (Embryophyta) entwickelten, hervor. Nu, Nucleus; Mi, Mitochondrium; EMS, Endomembransystem; PG, Peptidoglycan; IEM/OEM innere/äußere Hüllmembran (*inner/outer envelope membrane*); MP, Muroplast (Cyanell); RP, Rhodoplast; CP, Chloroplast. Modifiziert nach Moog (2009) und Gould et al. (2008).

Alle dieser höchst diversen Organismengruppen, welche auch als Archaeplastida bezeichnet werden (Adl et al, 2005), besitzen somit Plastiden mit zwei Hüllmembranen, die sich von einem cyanobakteriellen Vorläufer ableiten. Phylogenetische Analysen bekräftigen, dass es sich bei der primären Endosymbiose um ein monophyletisches Ereignis handelte (Chan et al, 2011; Martin et al, 2002; Price et al, 2012; Rodriguez-Ezpeleta et al, 2005; Yoon et al, 2004).

1.2.2 Eukaryote Endosymbiose - die Entstehung komplexer Plastiden

Neben den primären Plastiden der Archaeplastida, welche von zwei Membranen umgeben werden, finden sich in einer Vielzahl von Algen bzw. Protisten auch Plastiden, die von drei oder vier Hüllmembranen umschlossen sind (Cavalier-Smith, 2000). Diese werden aufgrund ihrer Morphologie als komplexe Plastiden bezeichnet und sind das Produkt der Internalisierung und Etablierung von in primärer Endosymbiose entstandener Organismen im Zuge weiterer Endosymbiosen höherer Ordnung (sekundäre Endosymbiose). Allerdings wurde dabei zumeist nicht nur die primäre Plastide an einen sekundären eukaryoten Wirt weitergegeben und beibehalten, sondern auch Elemente des primären Wirtes. Im Gegensatz zur monophyletischen primären Endosymbiose ereigneten sich mindestens drei voneinander unabhängige sekundäre Endosymbiosen (polyphyletisch), bei denen es sowohl zur Internalisierung einer Grünalge (zweimal) als auch Rotalge (mindestens einmal) durch einen eukaryoten Wirt kam (Gould et al, 2008; Keeling, 2009). Die Euglenophyta und Chlorarachniophyta, welche komplexe Plastiden mit drei bzw. vier Hüllmembranen besitzen, sind den Excavata bzw. Rhizaria zuzuordnen und evolvierten durch zwei unabhängige Endosymbioseereignisse mit einer Chlorophyte (Rogers et al, 2007). Die Linien der Cryptophyta, Haptophyta, Heterokontophyta (Chromista), sowie der Apicomplexa und Peridinin-enthaltenden Dinophyta (Alveolata) besitzen hingegen komplexe Plastiden rhodophytischen Ursprungs mit drei (Dinophyta) oder vier (alle anderen Linien) Hüllmembranen (Abb. 1.2) und werden unter dem Superphylum „Chromalveolata“ zusammengefasst (Cavalier-Smith, 1999).

Neben ökologisch höchst relevanten photosynthetischen Algengruppen, welche den Hauptanteil marinen Phytoplanktons bilden, werden den „Chromalveolaten“ auch heterotrophe Protisten zugerechnet. Apicomplexa besitzen, mit Ausnahme der Gattung *Cryptosporidium*, zwar komplexe Plastiden (Apicoplast), verloren jedoch die Kapazität des Organells zur Photosynthese und erhalten dieses für essentielle Stoffwechselwege, wie

Fettsäure-, Isoprenoid-, Eisen-Schwefel-Cluster- und Häm-Synthese (Woehle et al, 2011). Außerdem finden sich, neben *Cryptosporidium*, weitere plastidenlose Vertreter unter den Alveolaten (Ciliaten und einige Dinoflagellaten), sowie in den Linien der Cryptophyta und Heterokontophyta (Oomyceten), deren Genome z.T. anhand der Präsenz von vereinzelten Genen cyanobakterieller bzw. plastidärer Herkunft auf eine potentiell photosynthetische Vergangenheit hindeuten (Gould et al, 2008). Ob diese Gene jedoch das Produkt endosymbiontischen (EGT) oder lateralen Gentransfers (LGT) darstellen ist nicht eindeutig geklärt (Archibald, 2009).

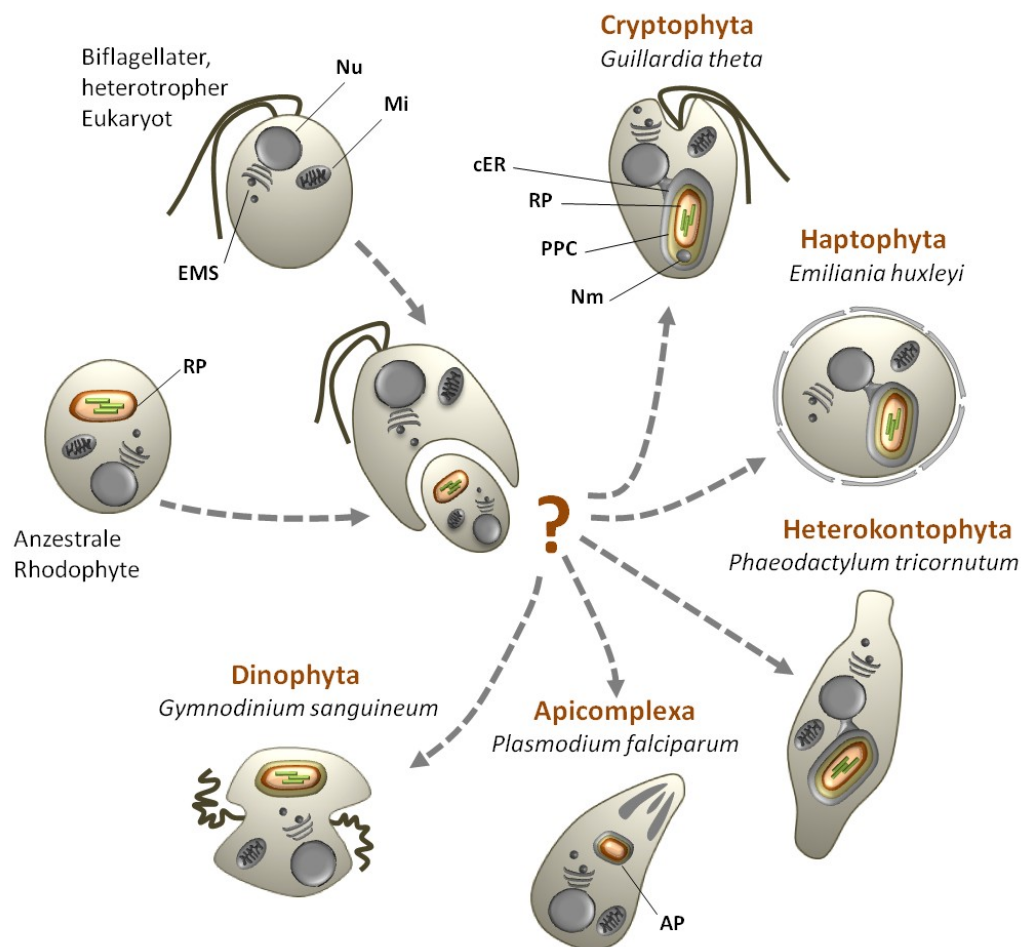


Abb. 1.2 Die Entstehung von Organismen mit komplexen Plastiden rhodophytischen Ursprungs. Gemäß der Chromalveolaten-Hypothese (Cavalier-Smith, 1999) geht die Entstehung von Organismen mit komplexen Plastiden rhodophytischen Ursprungs auf ein monophyletisches Ereignis sekundärer Endosymbiose zurück, bei dem eine anzestrale Rhodophyte von einem Eukaryoten Wirt internalisiert und als komplexe Plastide etabliert wurde. Aktuelle phylogenetische Analysen sprechen gegen eine Monophylie und deuten auf wesentlich komplexere evolutionäre Szenarien hin (siehe Text). Nu, Nucleus; Mi, Mitochondrium; EMS, Endomembransystem; RP, Rhodoplast; cER, *chloroplast endoplasmic reticulum*; PPC, periplastidäres Kompartiment (*periplastidal compartment*); Nm, Nucleomorph. Modifiziert nach Moog (2009) und Gould et al. (2008).

1999 wurde von Cavalier-Smith mit der Chromalveolaten-Hypothese ein monophyletischer Ursprung für Organismen mit komplexen Plastiden, die auf eine Rhodophyte zurückgehen, postuliert (Cavalier-Smith, 1999) (Abb. 1.2). Diese Hypothese stützte sich auf Ähnlichkeiten hinsichtlich Pigmentzusammensetzung und Symbiontenmorphologie dieser Linien. Ein wesentlicher Aspekt dieser Theorie setzte jedoch das Prinzip der evolutionären Sparsamkeit voraus, nämlich, dass ein solch komplexer Prozess wie die Etablierung eines eukaryoten Endosymbionten als Organell in Anbetracht der damit verbundenen notwendigen Regulationsmechanismen (Gentransfer, Proteinreimport, metabolische Kommunikation) in mehrfacher Entstehung unwahrscheinlich sei (Cavalier-Smith, 1999). Während phylogenetische Analysen die Chromalveolatenhypothese (Hackett et al, 2007; Harper & Keeling, 2003; Harper et al, 2005; Patron et al, 2007) oder zumindest die Monophylie der Plastiden dieser Organismen stützen (Bachvaroff et al, 2005; Fast et al, 2001; Patron et al, 2004; Yoon et al, 2002), werfen aktuelle Studien jedoch Zweifel hinsichtlich der Monophylie der „Chromalveolaten“ auf und deuten auf komplexere evolutionäre Szenarien hin.

Phylogenetische Analysen mit nucleuskodierten Proteinen von „Chromalveolaten“ weisen in diesem Kontext auf die Involvierung divergenter Wirtslinien in mehreren unabhängig stattgefundenen sekundären Endosymbioseereignissen hin (Burki et al, 2012b; Burki et al, 2007), oder schlagen eine initiale sekundäre Endosymbiose gefolgt von weiteren tertiären Endosymbioseereignissen vor (Baurain et al, 2010; Bodyl et al, 2009; Sanchez Puerta & Delwiche, 2008). Hinsichtlich Letzterem wurde postuliert, dass Cryptophyten und Haptophyten einen gemeinsamen sekundär-endocytobiotischen Ursprung besitzen, wobei Heterokontophyten und Alveolaten (Peridinin-enthaltende Dinoflagellaten und Apicomplexa) durch eine oder mehrere tertiäre Endosymbiosen mit einem Vorläufer aus dieser Linie entstanden sind (Archibald, 2009; Sanchez Puerta & Delwiche, 2008). Allerdings sprechen neuere phylogenetische Analysen wiederum gegen eine gemeinsame Wirtslinie der Cryptophyten und Haptophyten (Burki et al, 2012b). Letztere konnten dahingegen als Schwestergruppe einer als monophyletisch postulierten SAR-Klade, welche die Stramenopilen (Heterokontophyta), Alveolaten und Rhizaria zusammenfasst (Burki et al, 2007), identifiziert werden (Burki et al, 2012b).

Zunehmend komplexer wird die Situation durch die Identifizierung eines signifikanten Anteils von nucleuskodierten Genen mit vermeintlich chlorophytischem Ursprung in Diatomeen (Heterokontophyta), die auf einen kryptischen Grünalgenendosymbionten in diesen Organismen hindeutet, der möglicherweise verloren ging bevor es zur Akquirierung

des rhodophytischen Endosymbionten kam (Moustafa et al, 2009). Ein ähnliches Phänomen konnte auch für einen basalen photosynthetischen Alveolaten (*Chromera velia*) festgestellt werden (Woehle et al, 2011). Allerdings ist die Detektion eines „grünen Signals“ in den Genomen dieser Linien umstritten und kann, neben einer kryptischen Endosymbiose mit einer Chlorophyte, auch durch die Limitierung an verfügbaren Rhodophyten-Genomen und daraus resultierenden phylogenetischen Artefakten erklärt werden (Burki et al, 2012a; Moustafa et al, 2009; Woehle et al, 2011). Zudem kommt auch eine Akquirierung von „grünen Genen“ über lateralen Gentransfer in Frage (Dorrell & Smith, 2011).

Aktuell unterliegen Modelle zur Evolution von Organismen mit komplexen Plastiden rhodophytischen Ursprungs mit stetig wachsenden genomischen Daten einem ständigen Wandel. Die Prozesse die zur Entstehung dieser Organismengruppen geführt haben sind somit im Detail bislang unklar. Der überwiegende Teil phylogenetischer Analysen spricht jedoch gegen eine Entstehung von Organismen mit komplexen Plastiden rhodophytischer Herkunft über ein singuläres Ereignis sekundärer Endosymbiose.

Ähnlich der primären Endosymbiose ging die Etablierung sekundärer oder komplexer Plastiden mit einer drastischen genetischen Reorganisation einher, bei der redundante Symbiontengene verloren gingen und essentielle genetische Informationen in den Nucleus der Wirtszelle transferiert wurden (Kleine et al, 2009). Dabei erfolgte die Akquirierung genetischen Materials durch den Wirt nicht nur aus dem Plastidengenom, sondern vor allem aus dem Nucleus des eukaryoten Endosymbionten. Zusätzlich zur genetischen Verringerung, ereignete sich auch eine morphologische Reduktion des Endosymbionten, was bei Betrachtung der komplexen Plastiden sekundär evolvierter Algenlinien, die auf einen rhodophytischen Endosymbionten zurückgehen, offenbar wird (Abb. 1.2).

Die komplexen Plastiden der Cryptophyta, Haptophyta und Heterokontophyta (vier Hüllmembranen) befinden sich innerhalb einer Zisterne des rauen Endoplasmatischen Reticulums (ER) der Wirtszelle (Gibbs, 1979). Es wird vermutet, dass diese, auch als *chloroplast ER* bezeichnete Membran, aus der Fusion der phagotrophen Vakuole (in der sich der Endosymbiont befand) und dem ER der Wirtszelle hervorging und somit die äußerste Membran der komplexen Plastiden (cER-Membran) bildete (Cavalier-Smith, 1999). Bei den Apicomplexa besteht kein Kontinuum zwischen äußerster Hüllmembran der komplexen Plastide (epiplastidäre Membran) und ER, sodass der Apicoplast höchstwahrscheinlich frei im

Cytoplasma der Wirtszelle vorliegt (Sheiner & Striepen, 2012). Eine ähnliche Situation findet sich bei den Chlorarachniophyta.

Die zweitäußerste Membran komplexer Plastiden mit vier Hüllmembranen (PPM, periplastidäre Membran) wird von der ehemaligen Plasmamembran des Endosymbionten gebildet. Diese umschließt das periplastidäre Kompartiment (*periplastidal compartment*, PPC), welches auf das ursprüngliche Cytoplasma der internalisierten Archaeplastide-tragenden Alge zurückgeht. Dieses wurde zumeist drastisch reduziert, was, ausgenommen der primären Plastide, mit dem Verlust jeglicher Organellen des eukaryoten Endosymbionten einherging. Eine Ausnahme dabei bilden die Cryptophyten und Chlorarachniophyten, die jeweils einen remanenten, vom Endosymbionten (Rot- bzw. Grünalge) stammenden Nucleus, welcher als Nucleomorph bezeichnet wird, im PPC ihrer komplexen Plastiden erhalten (Moore & Archibald, 2009). Diese Situation wird als evolutionäres Intermediärstadium in Bezug auf die Reduktion der Endosymbionten angesehen (Maier et al, 2000). Die frei im Wirtscytoplasma vorliegenden komplexen Plastiden der Peridinin-enthaltenden Dinoflagellaten (Dinophyta) und Euglenophyten offenbaren den höchsten Grad der Symbiontenreduktion, da diese, im Gegensatz zu den zuvor beschriebenen Linien, nur drei Hüllmembranen besitzen (Green, 2011). Es wird postuliert, dass in beiden Linien die ehemalige Plasmamembran des jeweiligen Endosymbionten (PPM) vollständig verloren ging (Cavalier-Smith, 1999; Cavalier-Smith, 2000). Das innere Membranpaar (OEM/IEM, *outer/inner envelope membrane*) entspricht bei allen komplexen Plastiden höchstwahrscheinlich den beiden Hüllmembranen der ursprünglichen Archaeplastide des internalisierten eukaryoten Endosymbionten.

1.2. Proteinimport in primäre und komplexe Plastiden

Mechanismen endosymbiontischen Gentransfers sind dafür verantwortlich, dass eine Vielzahl von Plastidengenomen rezenter Organismen für weniger als 100 Proteine kodieren. Das bedeutet, dass ca. 95% der Gene für funktionelle Plastidenproteine im Genom der Wirtszelle lokalisiert sind und die Genprodukte in das Organell re-importiert werden müssen (Shi & Theg, 2012). Ein kritischer Schritt in der Evolution phototropher Eukaryoten war es daher, Transportmechanismen zu entwickeln, die eine stetige Versorgung der Plastiden mit solchen Proteinen von Seiten des Wirts gewährleisten. Diese müssen im Falle primärer Plastiden bis

zu drei und bei sekundären bzw. komplexen Plastiden sogar bis zu fünf biologische Membranen überqueren.

1.2.1 Proteinimport in primäre Plastiden

Die Synthese nucleuskodierter plastidärer Proteine erfolgt in Organismen mit primären Plastiden im Cytosol der Wirtszelle. Jene sind auf genetischer Ebene zumeist mit einer spezifischen plastidären Zielsteuerungssequenz, dem Transitpeptid, ausgestattet, welche sich auf Proteinebene als N-terminale Extension vor dem maturen Protein ausprägt. Obwohl Transitpeptide keine konservierte Primärsequenz besitzen, lassen sich einige gemeinschaftliche Charakteristika feststellen. So finden sich z.B. bei den Transitpeptiden der Archaeplastida gehäufte Vorkommen hydroxylierter Aminosäuren (AS) wie Serin und Threonin, ein erhöhtes Alanin-Level, sowie eine positive Nettoladung (Patron & Waller, 2007).

In höheren Pflanzen sind die Importprozesse, die zur Translokation von Präproteinen in primäre Plastiden führen, am besten studiert. Nach erfolgter Translation an cytosolischen 80S-Ribosomen werden plastidäre Präproteine von einem spezifischen *guidance complex* (Hsp70, 14-3-3) gebunden und zur äußeren Plastidenmembran geleitet (May & Soll, 2000). Der Transport über die beiden Plastidenhüllmembranen erfolgt über eine Maschinerie, die sich aus den multimeren TOC/TIC-Translokatoren (*translocon of the outer/inner chloroplast envelope membrane*) zusammensetzt. Zunächst erfolgt die Erkennung des Transitpeptids durch die membranständigen GTPasen Toc159 und Toc34 an der OEM (Jarvis, 2008). Ein weiterer importrelevanter Faktor des TOC-Komplex ist das Membranprotein Toc64, welches als Bindestelle für Hsp90 Chaperon-geleitete Imports substrate gilt (Sommer et al, 2012). Die Präproteine werden im Folgenden an die Translokationspore Toc75, der zentralen Komponente des TOC-Komplexes, herangeführt und über die OEM transloziert. Im Intermembranraum werden diese von einem Komplex aus Toc64, Toc 12, Hsp70, sowie dem verbindenden Element der TOC/TIC-Translokatoren, Tic22, weiter zum TIC-Komplex in der inneren Plastidenmembran (IEM) geleitet (Strittmatter et al, 2010). Kanalbildende Faktoren des TIC-Komplexes sind Tic20 Tic21 und Tic110, außerdem spielen einige Redox-regulatorische Komponenten eine Rolle (Tic32, Tic55, Tic62) (Soll & Schleiff, 2004). Eine weitere grundlegende Funktion wird von Tic40 erfüllt, welches ähnlich wie Tic110 als Bindestelle für stromale Chaperone gilt (Balseira et al, 2009). Letztere extrahieren das

Importsubstrat nach Eintritt in die TIC-Translokationspore in einem ATP-abhängigen Prozess ins Stroma, wo schließlich das Transitpeptid durch eine prozessierende Protease (SPP, *stromal processing protease*) abgespalten wird und die Proteine ihre native Konformation einnehmen (Shi & Theg, 2012). Bei Proteinen mit Funktionen in den Thylakoiden (z.B. Komponenten des Photosyntheseapparates) erfolgt die subsequente Translokation lumenaler Faktoren über Sec- oder Tat-Transportsysteme bzw. im Falle von Membranproteinen durch SRP-vermittelte oder spontane Integration (Albiniak et al, 2012).

Phylogenetische Analysen offenbaren, dass es sich bei einem signifikanten Teil des TOC/TIC-Translokators um Faktoren eukaryoter Herkunft handelt, während die übrigen (z.B. Toc75) einen cyanobakteriellen Ursprung aufweisen (Reumann et al, 2005; Reyes-Prieto et al, 2007). Generell finden sich die zentralen Komponenten der TOC/TIC-Translokatoren in allen Archaeplastiden tragenden Linien (Kalanon & McFadden, 2008; Price et al, 2012; Reumann et al, 2005; Steiner et al, 2005). Neben dem konventionellen TOC/TIC-vermittelten Importprozess liegen in höheren Pflanzen auch Hinweise auf einen vesikulären Transportmechanismus glycosylierter Proteinspezies zwischen ER, Golgi und primärer Plastide vor (Radhamony & Theg, 2006; Villarejo et al, 2005)

1.2.2 Proteinimport in komplexe Plastiden rhodophytischen Ursprungs

Der Import nucleuskodierter Präproteine in komplexe Plastiden rhodophytischen Ursprungs unterliegt im Vergleich zu primären Plastiden wesentlich aufwendigeren Mechanismen, da zum einen ein oder zwei zusätzliche Membranen passiert werden müssen und zum anderen, unabhängig von der Anzahl der Hüllmembranen, das Endomembransystem der Wirtszelle in die Transportvorgänge involviert ist. Die Majorität nucleuskodierter plastidärer Präproteine dieser Organismen ist mit einer, als BTS (*bipartite topogenic signal*) bezeichneten, zweigeteilten N-terminalen Zielsteuerungssequenz versehen (Cavalier-Smith, 2003). Diese setzt sich aus einem klassischen ER-Signalpeptid gefolgt von einer, den Transitpeptiden der Archaeplastida ähnelnden, TPL (*transit peptide-like sequence*) als N-terminale Extension vor Beginn des maturen Proteins zusammen (Gould et al, 2006a). Das Signalpeptid vermittelt den initialen, co-translationalen Transportschritt über Sec61 in der Membran des rauen ER (rER) der Wirtszelle (Bolte et al, 2009). In den Linien der Cryptophyten, Heterokontophyten und Haptophyten stellt dies bereits die Passage der ersten Hüllmembran der komplexen Plastide (4 Hüllmembranen) dar, da sich diese in Kontinuum mit dem rER befindet (cER). Es wird

vermutet, dass das SP nach erfolgter Präprotein-Translokation ins cER-Lumen prozessiert wird, wodurch im Weiteren die TPL exponiert wird (Apt et al, 2002). Die Signalfunktion der TPL ist ab diesem Punkt entscheidend, da verhindert werden muss, dass plastidäre Präproteine im cER den sekretorischen Weg einschlagen. Bei den Apicomplexa und Peridinin-enhaltenen Dinoflagellaten (Dinophyta), deren komplexe Plastiden von vier bzw. drei Membranen umgeben werden, erfolgt der Transport vom rER zur äußeren, epiplastidären Membran über einen vesikulären Mechanismus. Während dieser Transportweg bei den Dinophyta wahrscheinlich die Passage des Golgi Apparates miteinschließt, gelangen plastidäre Präproteine bei den Apicomplexa vermutlich auf direktem Wege von rER zum Apicoplasten (Patron et al, 2005; Sheiner & Striepen, 2012).

Es ist mittlerweile weitestgehend anerkannt, dass die Passage der zweitäußersten Membran der komplexen Plastide (PPM), welche die ehemalige Plasmamembran des rhodophytischen Endosymbionten darstellt, über das SELMA-System erfolgt (Agrawal et al, 2009; Felsner et al, 2011; Sheiner & Striepen, 2012). SELMA steht für *symbiont-specific ERAD-like machinery* und bezeichnet einen Translokationsapparat, der als ein vom ehemaligen ER des Rotalgenendosymbionten abgeleitetes und in die PPM der komplexen Plastide relokalisiertes ERAD-System gilt (ER-assoziierte Degradation) (Hempel et al, 2009; Sommer et al, 2007). Das ERAD-System eukaryoter Organismen dient für gewöhnlich der Erkennung und Dislokation fehlerhafter Proteine am ER. Diese werden in einem Ubiquitin-abhängigen Prozess aus dem ER ins Cytosol transloziert und anschließend proteasomal degradiert (Meusser et al, 2005). SELMA existiert allerdings parallel zum ERAD-System der Wirtszelle und wurde höchstwahrscheinlich im Laufe der Koevolution von Wirt und Symbiont spezifisch an den Import plastidärer Proteine angepasst (Bolte et al, 2011).

Erste Hinweise auf die Präsenz von ERAD-ähnlichen Faktoren in komplexen Plastiden rhodophytischen Ursprungs kamen aus der Cryptophyte *Guillardia theta*, deren Nucleomorph im PPC für einige ERAD-Kernkomponenten kodiert (Sommer et al, 2007). In nachfolgenden Studien konnte gezeigt werden, dass symbiontische ERAD- bzw. SELMA-Faktoren in den Genomen aller untersuchten Organismen mit komplexen Plastiden rhodophytischen Ursprungs, mit Ausnahme der Dinophyta (wahrscheinlich Verlust der PPM), konserviert sind und dort im PPC bzw. der PPM lokalisieren (Agrawal et al, 2009; Felsner et al, 2011; Hempel et al, 2009; Kalanon et al, 2009; Sommer et al, 2007; Spork et al, 2009). Unlängst gelang auch der experimentelle Nachweis der Involvierung des SELMA-Systems in Proteinimportprozesse über die PPM in den Apicomplexa. Es konnte gezeigt werden, dass der Import von Proteinen

in den Apicoplasten durch eine konditionale Herunterregulierung des als Translokationskanal über die PPM diskutierten sDer1 (s, symbionspezifisch) zum Erliegen kommt (Agrawal et al, 2009).

In der Heterokontophyte *Phaeodactylum tricornutum* ist das SELMA-System bisher am besten untersucht (Abb. 1.3). Nach dem von Sommer et al. (2007) postulierten Importmodell gelangen plastidäre Proteine durch sDer1-1/1-2 über die PPM ins PPC komplexer Plastiden rhodophytischen Ursprungs mit vier Hüllmembranen (Sommer et al, 2007). In diesem Kontext konnte von Hempel et al. (2009) in *P. tricornutum* gezeigt werden, dass diese PPM-Proteine in der Lage sind Homo- bzw. Heterooligomere auszubilden und mit TPL-Sequenzen von PPC-residenten, jedoch nicht von stromal lokalisierten Proteinen interagieren (Hempel et al, 2009). Es wird vermutet, dass die Translokation über die PPM in einem Ubiquitin (sUb)-abhängigen Prozess, ähnlich zu ERAD, abläuft, bei dem Substrate über eine typische Ubiquitinierungskaskade (Ub-aktivierendes E1 sUba1, Ub-konjugierendes E2 sUbc, Ub- E3-Ligase ptE3P) modifiziert werden und durch die AAA-ATPase sCdc48 samt Kofaktor sUfd1 ATP-abhängig ins PPC transloziert werden (Sheiner & Striepen, 2012; Sommer et al, 2007) (Abb. 1.3). Während es sich bei der potentiellen symbiontischen Ubiquitin-E3-Ligase bei Cryptophyten um ein Homolog zu dem aus ERAD bekannten Hrd1 handelt (Sommer et al, 2007), findet sich in *P. tricornutum* ein potentielles Ortholog (ptE3P) in der PPM, welches *in vitro* in der Lage ist Substrate zu ubiquitinieren (Hempel et al, 2010). Es wird postuliert, dass SELMA-Substrate, im Gegensatz zu einer Polyubiquitinierung bei ERAD-Substraten, im Zuge des Importprozesses monoubiquitiniert werden, da dem symbiontischen Ubiquitin von *P. tricornutum* das zur Polyubiquitinierung benötigte Lysin an Position 48 fehlt (Sommer et al, 2007). Außerdem wird vermutet, dass potentiell ubiquitinierte SELMA-Substrate nach dem Import ins PPC wieder deubiquitiniert werden und nicht wie bei ERAD dem Proteasomkomplex zur Degradation zugeführt werden. In diesem Kontext konnte eine PPC-lokalisierte Deubiquitinase (ptDUP) in *P. tricornutum* identifiziert werden, welche *in vitro* dazu befähigt ist Ubiquitin von synthetischen Substraten abzuspalten (Hempel et al, 2010).

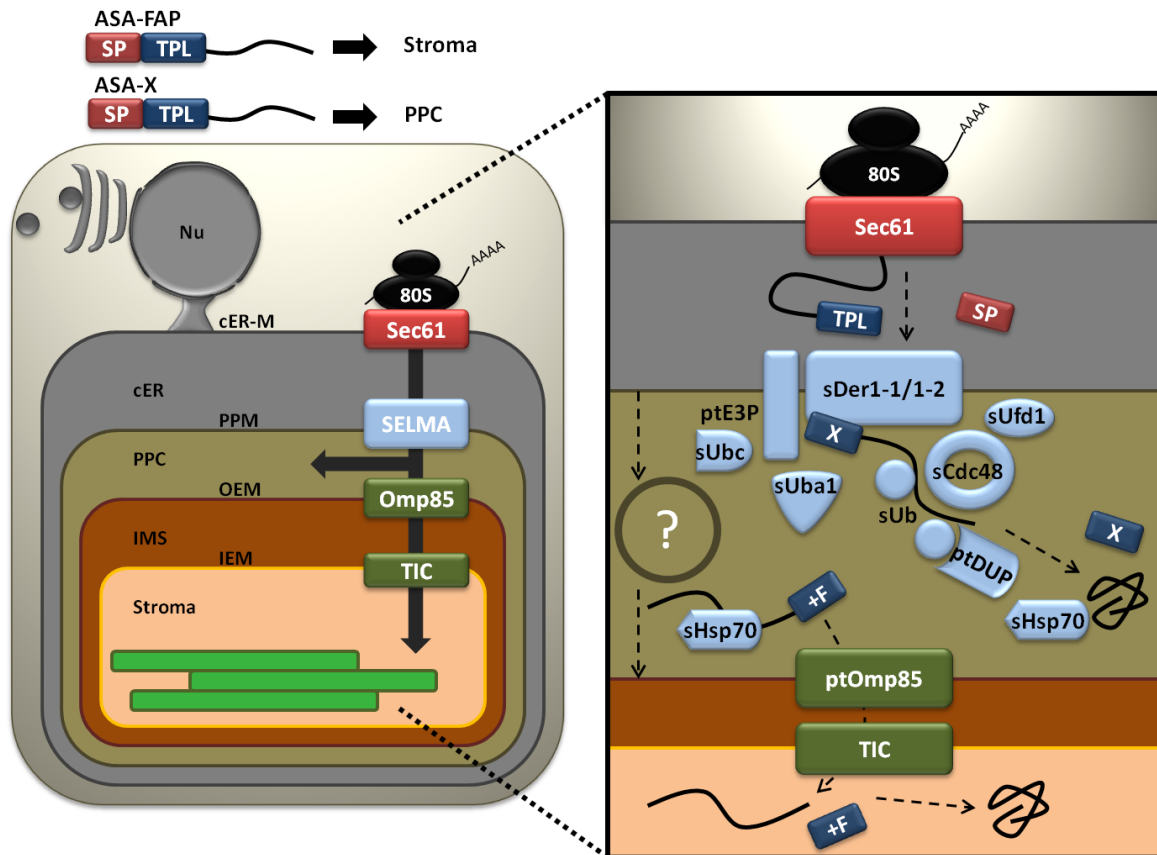


Abb. 1.3 Import nucleuskodierter Proteine in komplexe Plastiden rhodophytischen Ursprungs am Beispiel der Diatomee *P. tricornutum*. Nucleuskodierte plastidäre Proteine werden als Präproteine mit einer N-terminalen BTS (*bipartite topogenic signal*), bestehend aus SP (Signalpeptid) und TPL (*transit peptide-like sequence*), im Cytosol der Wirtszelle an 80S-Ribosomen synthetisiert. Das Signalpeptid vermittelt den co-translationalen Import ins cER über Sec61 und wird daraufhin wahrscheinlich prozessiert. Die Passage der zweiten Membran (PPM) erfolgt über das SELMA-System (*symbiont specific ERAD-like machinery*) höchstwahrscheinlich in einem Ubiquitin (sUb)-abhängigen Prozess. Dabei werden Präproteine im Zuge des Imports über sDer1-1/1-2 vermutlich ubiquitiniert (sUba1, sUbc, ptE3P), von sCdc48 samt Kofaktor sUfd1 ins PPC gezogen und anschließend deubiquitiniert (ptDUP). PPC-residente Proteine werden nach dem Import über SELMA durch einen unbekannten Faktor prozessiert und Chaperon-vermittelt (sHsp70) gefaltet. Stromale Proteine (konserviertes „ASA-FAP-Motiv“) werden im Weiteren wahrscheinlich über ptOmp85 (OEM) und einen TIC-Komplex (IEM) über die beiden inneren Membranen der komplexen Plastide transloziert, prozessiert und gefaltet. Alternativ wird ein vesikulärer Mechanismus zwischen PPM und OEM postuliert (Gibbs, 1979). cER(-M), *chloroplast ER(-Membran)*; PPM, *periplastidäre Membran*; PPC, *periplastidäres Kompartiment*; OEM, *plastid outer envelope membrane*; IMS, *intermembrane space*; IEM, *plastid inner envelope membrane*; Nu, Nucleus; +, positive TPL-Nettoladung; F, aromatische AS oder Leucin; X, weder aromatische AS noch Leucin. Abbildung nach Sommer et al. (2007), Hempel et al. (2009), Hempel et al. (2010), Bullmann et al. (2010).

Nach erfolgtem Präproteinimport in das PPC wird das TPL PPC-residenter Proteine vermutlich von einem bisher unbekannten Faktor prozessiert und die Proteine werden durch ein PPC-lokalisiertes sHsp70-Chaperon gefaltet, wobei stromale Präproteine von Letzterem

wahrscheinlich in einer transportkompetenten Form gehalten werden (Gould et al, 2006b; Sommer et al, 2007) (Abb. 1.3).

Vor und nach dem Import von Präproteinen ins PPC über SELMA ist die Signalfunktion der TPL entscheidend um eine korrekte Sortierung der Proteine innerhalb der Subkompartimente der komplexen Plastide zu gewährleisten. Ähnlich zu den Transitpeptiden der Rhodophyta, zeichnen sich plastidäre bzw. stromale Präproteine der meisten Organismen mit komplexen Plastiden rhodophytischen Ursprungs durch die Präsenz eines Phenylalanins (F) an der ersten Position der TPL aus (Patron & Waller, 2007). Dieses findet sich, vorwiegend in den Linien der Cryptophyten und Heterokontophyten, zumeist im Kontext eines spezifischen „ASA-FAP-Motivs“ am Übergang von SP zu TPL (Gould et al, 2006a; Gruber et al, 2007). Bei Haptophyten und Apicomplexen (sowie Dinophyten) scheint dieses Motiv weniger stark konserviert zu sein (Patron & Waller, 2007). Eine Reihe genomischer und experimenteller Analysen in *P. tricornutum* ergaben, dass neben dem dominant auftretenden F auch andere aromatische AS, wie Tyrosin (Y) und Tryptophan (W), oder sogar Leucin (L) diese Position einnehmen können um ein plastidäres *targeting* zu vermitteln (Gruber et al, 2007; Kilian & Kroth, 2005). Das entscheidende Charakteristikum von PPC-Proteinen ist dahingegen das Fehlen einer solchen aromatischen AS oder Leucin an der +1-Position der TPL (Gould et al, 2006a; Gruber et al, 2007) (Abb. 1.3). Neben der Natur der ersten AS der TPL hat die Nettoladung dieser Zielsteuerungssequenz einen Einfluss auf die Transportmechanismen ab der PPM. Felsner et al. (2010) konnten in *P. tricornutum* zeigen, dass eine diskrete positive TPL-Nettoladung eine Rolle beim Transport über die PPM spielen kann, während eine solche für die Passage der beiden inneren Plastidenmembranen (OEM/IEM) sogar notwendig zu sein scheint (Abb. 1.3). Außerdem konnte festgestellt werden, dass auch der N-Terminus des maturen Proteins, ergänzend zur TPL, eine Auswirkung auf das *targeting* von Präproteinen ausüben kann (Felsner et al, 2010).

Der Import plastidärer Proteine aus dem PPC ins Stroma verläuft bei komplexen Plastiden rhodophytischen Ursprungs höchstwahrscheinlich ähnlich zu den aus Archaeplastida bekannten Mechanismen ab. Allerdings scheinen typische TOC-Faktoren in der OEM dieser Organismen zu fehlen. Dahingegen konnte in *P. tricornutum* mit ptOmp85 (*outer membrane protein 85*) ein in der OEM lokalisiertes β -barrel Protein mit elektrophysiologischen Eigenschaften ähnlich zu Proteinen der Toc75- bzw. Omp85-Familie aus primären Plastiden bzw. Gram-negativen Bakterien identifiziert werden (Bullmann et al, 2010). Homologe zu ptOmp85 finden sich auch in Haptophyten und Apicomplexen. Vorläufige experimentelle

Resultate aus *Toxoplasma gondii* (Apicomplexa) deuten drauf hin, dass der Verlust von Omp85 in einem ausgeprägten Importdefekt apicoplastidärer Proteine resultiert (Sheiner & Striepen, 2012).

Für Organismen mit komplexen Plastiden rhodophytischen Ursprungs, welche von vier Membranen umgeben werden, wurde bereits 1979 ein vesikuläres Transportmodell für plastidäre Präproteine über das PPC postuliert (Gibbs, 1979). Nach dieser Hypothese gelangen Proteine, die für die primäre Plastide bestimmt sind, von der PPM in Vesikeln über ein periplastidäres Reticulum im PPC zur äußeren Membran der primären Plastide (OEM) (Gibbs, 1979) (Abb. 1.3). Diese Hypothese nach Gibbs beruhte im Wesentlichen auf der Beobachtung vesikulärer Strukturen im PPC im Rahmen von elektronenmikroskopischen Studien, konnte jedoch bisher experimentell weder verifiziert noch widerlegt werden. Allerdings existieren, wie zuvor dargelegt, mit SELMA und Omp85 Evidenzen für alternative Transportprozesse über PPM und OEM.

Hinsichtlich der Passage der inneren Membran komplexer Plastiden rhodophytischer Herkunft liegen Hinweise für die Präsenz einer Reihe von TIC-Komponenten in allen bisher untersuchten Linien (Cryptophyta, Heterokontophyta, Apicomplexa) vor (McFadden & van Dooren, 2004). Experimentelle Evidenzen einer Involvierung dieser Faktoren in den Importprozess von Präproteinen über die IEM kommen erneut aus den Apicomplexa. In *T. gondii* konnte gezeigt werden, dass sowohl Tic20 als auch Tic22 eine essentielle Transportfunktion bei der Passage von apicoplastidären Präproteinen über die IEM wahrnehmen (Sheiner & Striepen, 2012; van Dooren et al, 2008).

1.3 Das periplastidäre Kompartiment – ein minimiertes eukaryotes Cytoplasma

Organismen mit komplexen Plastiden, welche von vier Membranen umgeben werden, verfügen über zwei evolutionär divergente eukaryote Cytoplasmen. Neben dem Cytosol der Wirtszelle existiert in ihren komplexen Plastiden mit dem PPC (periplastidäres Kompartiment) ein weiteres, jedoch stark reduziertes Cytoplasma, das auf das ehemalige Cytoplasma des internalisierten eukaryoten Endosymbionten zurückgeht (Cavalier-Smith, 2003). Während dieses Kompartiment in den meisten Linien genetisch und morphologisch drastisch reduziert wurde (Heterokontophyta, Haptophyta, Apicomplexa), oder

wahrscheinlich sogar verloren ging (Dinophyta), unterhalten die Cryptophyta, wie bereits erwähnt (1.2.2), einen Restzellkern des rhodophytischen Endosymbionten (Nucleomorph) und somit ein genetisch aktives PPC. Die Sequenzierung einiger Nucleomorph-Genome (*Guillardia theta*, *Hemiselmis andersenii*, *Cryptomonas paramecium*, *Chroomonas mesostigmatica*) ermöglicht Einblicke in die im PPC ablaufenden Prozesse dieser Organismen (Douglas et al, 2001; Lane et al, 2007; Moore et al, 2012; Tanifuji et al, 2011).

Die bisher sequenzierten Nucleomorph-Genome der Cryptophyten weisen eine genetische Kapazität von 486 bis 703 Kbp (Kilobasenpaare) auf und bestehen jeweils aus drei winzigen, linearen Chromosomen mit z.T. variierender Komposition (Moore et al, 2012). Sie kodieren zwischen 518 und 780 Gene, deren Produkte vorwiegend in *housekeeping*-Funktionen (Transkription, Translation, DNA- und RNA-Metabolismus, Zellzyklus-Kontrolle und Mitose, sowie Proteinfaltung und Degradation) involviert sind (Douglas et al, 2001; Lane et al, 2007; Moore et al, 2012; Tanifuji et al, 2011). 216 dieser Gene finden sich konserviert auf allen vier sequenzierten Nucleomorph-Genomen wieder. Neben Genen, die für bekannte Funktionen kodieren, findet sich auch ein signifikanter Anteil (ca. 40%) an ORFs (*open reading frames*), denen bisher keine Funktionen zugeordnet werden konnten. Bemerkenswert ist, dass die Nucleomorph-Genome von *G. theta*, *H. andersenii* und *C. mesostigmatica* ein identisches Set von 31 essentiellen plastidären Proteinen kodieren (Moore et al, 2012), deren Präsenz als ein Grund für die Erhaltung des remanenten Rotalgennucleus gilt (Douglas et al, 2001). Bei *C. paramecium* handelt es sich um eine nicht-photosynthetische Cryptophyte. Dennoch kodiert das Nucleomorph-Genom dieser Alge für 18 plastidär-assoziierte Proteine, von denen sich 13 mit denen anderer Cryptophyten decken (Tanifuji et al, 2011).

Abseits von Funktionen der genetischen Selbsterhaltung und plastidären Proteinen, finden sich in allen vier bekannten Nucleomorph-Genomen Gene für symbiontische ERAD-Faktoren (Douglas et al, 2001; Lane et al, 2007; Moore et al, 2012; Tanifuji et al, 2011), welche der SELMA-Translokationsmaschinerie für plastidäre Präproteine über die PPM zugerechnet werden (Sommer et al, 2007) (siehe 1.2.2). Für Proteine mit Beteiligung an metabolischen Synthesewegen kodierende Gene konnten, mit Ausnahme einer Geranyl-Geranyl-Transferase (Carotinoid-Biosynthese) nicht auf den Nukleomorph-Genomen identifiziert werden. Allerdings ist bekannt, dass bei Cryptophyten die Synthese und Speicherung von Stärke ausschließlich im PPC abläuft (Deschamps et al, 2006). Beteiligte Enzyme sind daher höchstwahrscheinlich im Genom der Wirtszelle kodiert und müssen aus dem Cytosol ins PPC re-importiert werden, was bereits für eine UGG-Transferase (*UDP-glucose – starch glycosyl*

transferase) in heterologen Lokalisationsstudien gezeigt werden konnte (Gould et al, 2006a). Wie viele PPC-Proteine in Cryptophyten letztendlich nucleuskodiert vorliegen und welche weiteren Prozesse im PPC ablaufen, ist bisher aufgrund fehlender Genomdaten unbekannt. Allerdings steht die Publikation der Nucleus-Genomsequenz von *G. theta* kurz bevor und könnte hier entscheidende Hinweise liefern. Die Präsenz essentieller plastidärer Gene und von Genen kodierend für SELMA-Faktoren auf den Nucleomorph-Genomen, sowie der PPC-ansässige Stärke-Metabolismus, gelten jedoch als bedeutende Gründe für den Erhalt des Nucleomorphs bzw. PPCs in Cryptophyten (Douglas et al, 2001).

Im PPC der Heterokontophyten findet, im Gegensatz zu Cryptophyten, keine Genexpression statt, da jegliche genetische Information des Rotalgennucleus entweder verloren ging oder im Zuge endosymbiontischen Gentransfers in den Nucleus der Wirtszelle transferiert wurde. PPC-residente Proteine sind demnach im Genom der Wirtszelle kodiert und müssen, wie unter 1.2.2 dargestellt, ins PPC re-importiert werden. Ein weiterer Unterschied zu Cryptophyten ist, dass Heterokontophyten keine Stärke im PPC speichern, sondern die Energiespeicherung in Form von Lipiden oder Chrysolaminarin im Wirtscytosol bzw. der Vakuole abläuft (Kroth et al, 2008). Das PPC der Heterokontophyten ist darüber hinaus frei von Organellen, wie z. B. ER oder Mitochondrien (Keeling, 2010), was eine morphologische Kompaktierung erlaubt. In der Diatomee *P. tricornutum* weist das PPC eine sehr spezielle, charakteristische Morphologie auf (Abb 1.4). Während die beiden inneren Membranen der komplexen Plastide (OEM und IEM) das Stroma flügelartig mit einer Engstelle im Zentrum umgeben, weicht die PPM teilweise von dieser Formation ab. Insbesondere im Bereich der Engstelle von IEM und OEM entfernt sich die PPM von der OEM wodurch das PPC dort etwas an Raum gewinnt, wohingegen die PPM ansonsten der OEM sehr eng anliegt. Diese typische Morphologie lässt sich sowohl in elektronenmikroskopischen Analysen (Abb. 1.4) als auch bei *in vivo* Lokalisationsstudien beobachten, wo PPC-dirigiertes GFP in der im Zentrum der komplexen Plastide gelegenen, PPC-typischen „Blob-ähnlichen“ Fluoreszenzstruktur (BLS) resultiert (Kilian & Kroth, 2005) (siehe auch Abb. 1.6).

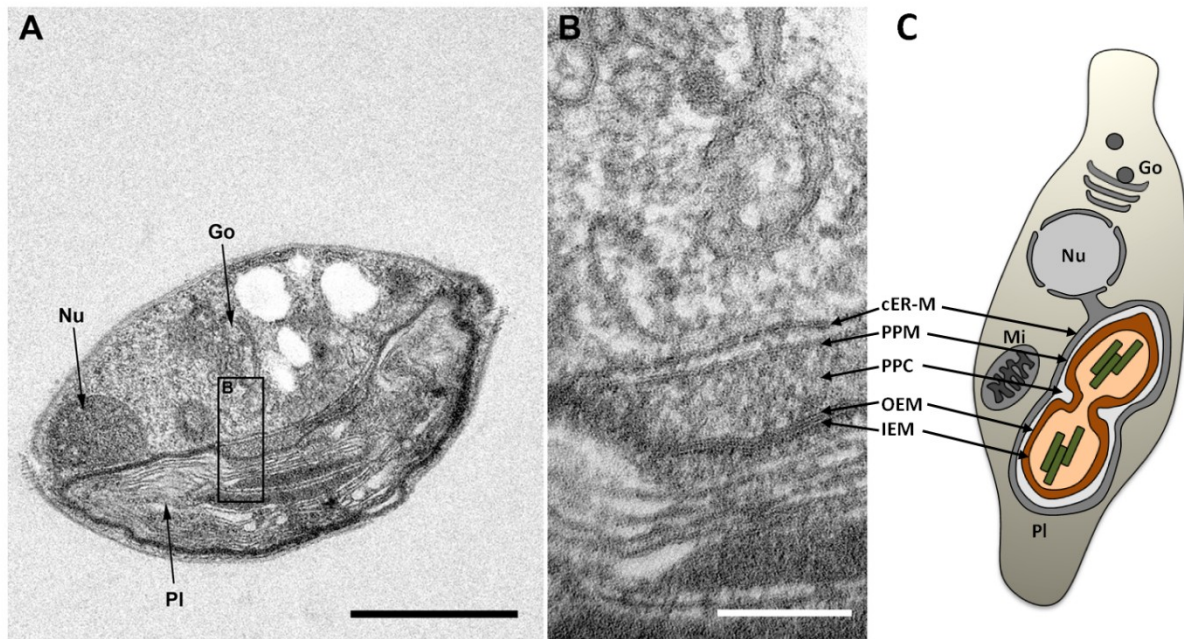


Abb. 1.4 Morphologie des periplastidären Kompartiments (PPC) von *P. tricornutum*. (A) Elektronenmikroskopische Aufnahme einer *P. tricornutum* Zelle. Maßstab 2 μm . (B) Vergrößerung des Rahmens aus A. Das periplastidäre Kompartiment befindet sich zwischen den beiden äußeren (cER-M/PPM) und inneren Membranen (OEM/IEM) der komplexen Plastide. OEM und IEM bilden eine Engstelle im Zentrum, wodurch das PPC dort an Raum gewinnt. Maßstab 200 nm. (C) Schematische Darstellung der Zellstruktur von *P. tricornutum*. cER(-M), *chloroplast* ER(-Membran); PPM, periplastidäre Membran; PPC, periplastidäres Kompartiment; OEM, *plastid outer envelope membrane*; IEM, *plastid inner envelope membrane*; Go, Golgi Apparat; Nu, Nucleus, Mi, Mitochondrium, Pl, Plastide. Abbildung A und B wurden von Dr. Andreas Klingl zu Verfügung gestellt und zeigen eine Polyhydroxybutyrat-synthetisierende *P. tricornutum* Zelle (Hempel et al, 2011).

Über die Funktionen und Gründe für den Erhalt des PPC in Heterokontophyten ist bislang nur wenig bekannt. Allerdings gelang es bereits über genomische Analysen und *in vivo* Lokalisationsstudien mittels GFP-Fusionsproteinen einige wenige Proteine im PPC der Diatomee *P. tricornutum* zu identifizieren. Das derzeitige Wissen über das PPC-Proteom von *P. tricornutum* beläuft sich, wie unter 1.2.2 erwähnt, vorwiegend auf Komponenten der SELMA-Translokationsmaschinerie für nucleuskodierte Präproteine über die PPM der komplexen Plastide (Hempel et al, 2009; Sommer et al, 2007). Das SELMA-System in *P. tricornutum* setzt sich, soweit bisher bekannt, aus den folgenden Komponenten zusammen: den membranständigen, als Translokationspore diskutierten Proteinen sDer1-1 und sDer1-2; der Ubiquitinierungsmaschinerie, bestehend aus Ubiquitin (sUb), einem Ubiquitin-aktivierenden E1 Enzym (sUba1), zwei Ubiquitin-konjugierenden E2 Enzymen (sUbc4/6) und einer membranintegralen E3 Ubiquitin-Ligase (ptE3P); der zentralen energieliefernden AAA-ATPase sCdc48; dem Cdc48-Kofaktor sUfd1; sowie einem deubiquitinierenden Enzym (ptDUP) (Hempel et al, 2009; Hempel et al, 2010; Sommer et al, 2007) (Abb. 1.5). Außerdem

konnte mit sHsp70 ein PPC-lokalisiertes Chaperon in *P. tricornutum* identifiziert werden, welches dort wahrscheinlich an der Faltung PPC-residenter Proteine, sowie am Transport plastidärer Präproteine über die dritte Membran der komplexen Plastide (OEM) mitwirkt (Gould et al, 2006b; Sommer et al, 2007) (Abb. 1.5).

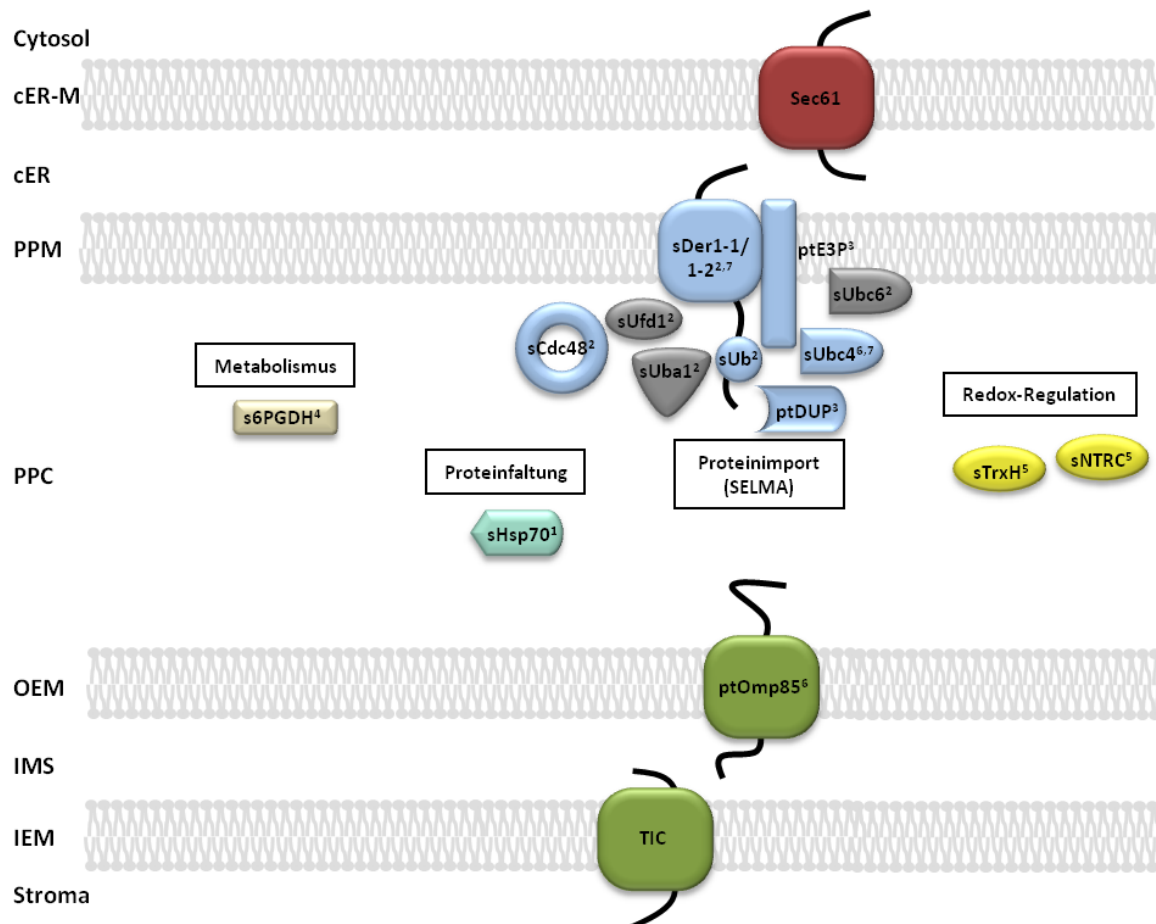


Abb. 1.5 Das PPC der Diatomee *Phaeodactylum tricornutum*. Resultierend aus bisherigen Studien finden sich im PPC von *P. tricornutum* putative Funktionen hinsichtlich Proteinimport (SELMA), Proteinfaltung, Redox-Regulation und Metabolismus. Faktoren, mit verifizierter PPC-Lokalisation (Präsequenz- oder volle Länge-GFP Lokalisationsstudien) sind farbig dargestellt und entsprechend gekennzeichnet. Proteine die bisher *in silico* als potentiell PPC-lokalisiert eingestuft wurden, sind in dunkelgrau dargestellt. ¹, Gould et al. (2006); ², Sommer et al. (2007); ³, Hempel et al. (2010); ⁴, Gruber et al. (2009); ⁵, Weber et al. (2009); ⁶, Bullmann et al. (2010); ⁷, Hempel et al. (2009). cER(-M), *chloroplast ER(-Membran)*; PPM, *periplastidäre Membran*; PPC, *periplastidäres Kompartiment*; OEM, *plastid outer envelope membrane*; IMS, *intermembrane space*; IEM, *plastid inner envelope membrane*.

Erste Hinweise auf ein Protein mit metabolischer Relevanz im PPC ergaben sich 2009 in einer Studie von Gruber et al. Über eine Kombination von *in silico* Analysen und *in vivo* Lokalisationsstudien von Präsequenzen identifizierter Faktoren des reduktiven und oxidativen

Pentosephosphatwegs fusioniert an GFP, gelang es dabei eine 6-Phosphogluconolacton-Dehydrogenase (s6PGDH) als PPC-lokalisiert in *P. tricornutum* nachzuweisen (Gruber et al, 2009) (Abb. 1.5). Dieses Enzym katalysiert die Umwandlung von Gluconat-6-Phosphat und CO₂ zu Ribulose-5-Phosphat, wobei NADPH als Reduktionsäquivalent entsteht. Dabei handelt es sich um einen Teilschritt des oxidativen Pentosephosphatwegs, der in Diatomeen primär im Cytosol abläuft (Gruber et al, 2009). Es wird vermutet, dass im PPC entstandenes Ribulose-5-Phosphat dem stromalen Calvin-Zyklus (reduktiver Pentosephosphatweg) zugeführt bzw. in der Plastide für anabolische Zwecke genutzt wird (Gruber et al, 2009).

Des Weiteren liegen Evidenzen für zwei Redox-regulatorische Enzyme im PPC von *P. tricornutum* vor. Dabei handelt es sich um sTrxH (Thioredoxin H) und eine NADPH-abhängige Thioredoxin-Reduktase (sNTRC, *NADPH depending thioredoxin reductase containing N-terminal thioredoxin domain*) (Weber et al, 2009) (Abb. 1.5). Thioredoxine sind dafür bekannt Disulfidbrücken über ein konserviertes WC(G/P)PC-Motiv an Zielproteinen zu generieren, welche reversibel von Thioredoxin-Reduktasen wieder aufgelöst werden können um spezifisch die Aktivität von Zielproteinen zu regulieren (Weber et al, 2009). Bezüglich möglicher Zielproteine für das Thioredoxin-System im PPC von *P. tricornutum* liegen jedoch bisher keine Hinweise vor.

Insgesamt bestehen somit bisher Evidenzen für 14 PPC- bzw. PPM-residente Proteine in *P. tricornutum*, wobei die Lokalisation von 11 dieser Faktoren durch Expression von GFP-Fusionsproteinen (BTS- oder volle Länge-GFP) verifiziert wurde (Abb. 1.5). Alle diese Proteine besitzen eine PPC-spezifische Zielsteuerungssequenz in Form einer BTS, bei der die erste Position der TPL nicht von einer aromatischen AS oder Leucin eingenommen wird. Hinsichtlich bisheriger Funktionen im PPC von *P. tricornutum* scheint der Proteinimport über das SELMA-System eine dominante Rolle zu spielen. Kernfaktoren von SELMA finden sich, wie bereits unter 1.2.2 erwähnt, auch im PPC aller anderen Organismen mit komplexen Plastiden rhodophytischen Ursprungs.

1.4. Triosephosphat/Phosphat-Translokatoren in primären und komplexen Plastiden

Die treibende Kraft für heterotrophe Eukaryoten zur Photosynthese befähigte Organellen (primäre und komplexe Plastiden) zu akquirieren (Endosymbiose) war der daraus

resultierende Vorteil photoautotropher Lebensweise. Dabei werden vom Symbionten synthetisierte energiereiche Metabolite zur Verfügung gestellt, welche von der Wirtszelle metabolisiert werden können (Energiespeicherung, Anabolismus) (Weber, 2004). Solche Metabolite werden in Plastiden phototropher Organismen u.a. im Rahmen des Calvin-Zyklus als Triosephosphate synthetisiert (Photosynthese, Fixierung von CO₂) (Flügge, 1999). Eine wichtige Voraussetzung zur erfolgreichen Etablierung des Endosymbionten als Organell im Rahmen der primären oder eukaryoten Endosymbiose war es daher den Metabolismus des Endosymbionten mit dem der Wirtszelle zu verbinden (Weber et al, 2006). Die Hüllmembranen des Endosymbionten bildeten dabei die limitierende Barriere. Während Metabolite bei primären Plastiden der Pflanzen, Chlorophyta, Rhodophyta und Glaucophyta zwei Membranen überqueren müssen, bei denen wahrscheinlich lediglich die innere Membran (IEM) selektiv passiert werden muss (Duy et al, 2007), müssen diese bei komplexen Plastiden sekundär evolvierter Organismen eine oder zwei zusätzliche Membranen überqueren um ins Cytosol der Wirtszelle zu gelangen.

Im Zuge der Evolution entwickelten sich bei den Archaeplastida Transportsysteme, die den Metabolitaustausch zwischen Symbiont und Wirt gewährleisten konnten. Es wird vermutet, dass dabei initial ein aus dem Endomembransystem der Wirtszelle stammender Nukleotidzuckertranslokator (NST, *nucleotide sugar translocator*) in die IEM integriert wurde (Colleoni et al, 2010). Von diesem lassen sich vermutlich alle heutigen plastidären Phosphat-Translokatoren ableiten (Colleoni et al, 2010; Weber et al, 2006). Letztere umfassen bei höheren Pflanzen Triosephosphat/Phosphat-Translokatoren (TPTs), Glucose-6-Phosphat/Phosphat-Translokatoren (GPT), Xylulose-5-Phosphat/Phosphat-Translokatoren (XPT), sowie Phosphoenolpyruvat/Phosphat-Translokatoren (PPT) (Weber & Linka, 2011). Bei diesen handelt es sich um Membranproteine, die funktionell als Dimere existieren und ihre Substrate über einen strikten Antiportprozess transportieren (Flügge, 1999). In Rotalgen konnte bisher nur die Präsenz plastidärer TPTs und PPTs nachgewiesen werden, wohingegen in den Glaucophyten typische plastidäre Phosphat-Translokatoren zu fehlen scheinen (Weber & Linka, 2011). Kürzlich gelang jedoch die Identifizierung von putativ plastidären Hexosephosphat-Translokatoren des bakteriellen UhpC-Typs in der Glaucophyte *Cyanoecista paradoxa*, zu denen sich auch Homologe in Grün- und Rotalgen finden (Price et al, 2012).

In Organismen mit komplexen Plastiden rhodophytischen Ursprungs finden sich ebenfalls plastidäre Phosphat-Translokatoren. Allerdings konnten in diesen weder GPTs, XPTs noch PPTs identifiziert werden, sondern ausschließlich Translokatoren des TPT-Typs, für die ein

monophyletischer Ursprung postuliert wird (Weber et al, 2006; Weber & Linka, 2011). In den Apicomplexa und Cryptophyta wurden plastidäre TPTs bereits näher untersucht. Apicomplexa besitzen, wie unter 1.1.2 dargestellt, eine komplexe Plastide ohne Befähigung zur Photosynthese. Dies setzt voraus, dass die essentiellen biochemischen Prozesse im Apicoplasten mit Kohlenstoffgerüsten der Wirtszelle versorgt werden, da diese aufgrund fehlender Photosynthesekapazität nicht im Apicoplasten synthetisiert werden können (Fleige et al, 2007). Sowohl *Plasmodium falciparum* als auch *Toxoplasma gondii* verfügen über TPTs in den Membranen ihres Apicoplasten (Fleige et al, 2007; Mullin et al, 2006). Während in *P. falciparum* mit PfoTPT und PfiTPT zwei TPTs existieren, die in der äußeren bzw. inneren Membran des Apicoplasten lokalisieren (Mullin et al, 2006), besitzt *T. gondii* nur einen apicoplastidären TPT, für den eine Lokalisation in multiplen Membranen des Apicoplasten vermutet wird (Fleige et al, 2007). Eine Analyse der TPTs über Proteoliposomen ergab, dass diese ähnliche Substratspektren besitzen und in der Lage sind Dihydroxyacetonphosphat (DHAP, Triosephosphat), Phosphoenolpyruvat (PEP), sowie 3-Phosphoglycerat (3PGA) im Austausch mit anorganischem Phosphat (Pi) zu transportieren, während phosphorylierte Hexosen (Glucose 6-Phosphat, G6P) nicht als Substrate akzeptiert werden (Brooks et al, 2010; Lim et al, 2010). Es wird vermutet, dass die Substrate apicoplastidärer TPTs, welche Zwischenprodukte der Glycolyse darstellen, im Apicoplasten als Grundgerüste für die Fettsäure- und Isoprenoidbiosynthese genutzt werden können (Brooks et al, 2010; Lim et al, 2010; Weber & Linka, 2011).

Weitere Einblicke zu TPTs in komplexen Plastiden rhodophytischer Herkunft kommen aus der Cryptophyte *G. theta*. Haferkamp et al. (2006) konnten über die Generierung von Proteoliposomen, entweder bestehend aus der Gesamtmembranfraktion oder den beiden inneren Plastidenmembranen von *G. theta*, nachweisen, dass sowohl ein Transport von DHAP, als auch PEP erfolgt, wobei 3PGA und G6P nicht transportiert werden (Haferkamp et al, 2006). Es wird postuliert, dass dieser Transport von mindestens zwei genomisch identifizierten TPTs (TPT1 und TPT2) vollzogen wird, von denen TPT2 möglicherweise in der IEM und TPT1 vermutlich in der cER-Membran oder PPM lokalisiert (Haferkamp et al, 2006). Transkriptionelle Analysen beider TPTs ergaben, dass diese lichtabhängig exprimiert werden. Während TPT1-mRNA hauptsächlich in der Dunkelphase bzw. bei Nacht akkumuliert, wird TPT2 vorwiegend bei Licht exprimiert, eine Situation die direkt mit dem Stärke-Metabolismus korreliert. Es wird vermutet, dass durch Photosynthese generierte Triosephosphate am Tag über TPT2 ins PPC transportiert und dort in Form von Stärke gespeichert werden, wobei Stärke-Degradationsprodukte bei Nacht wahrscheinlich über TPT1

ins Cytosol der Wirtszelle gelangen um dort für anabolische Zwecke genutzt werden (Haferkamp et al, 2006).

Hinsichtlich der Präsenz von TPTs in Peridinin-enhaltenden Dinoflagellaten (Dinophyta) liegen bisher aufgrund fehlender Genomdaten keine Evidenzen vor. In Heterokontophyten (und Haptophyten) existieren dahingegen genomische Hinweise für die Existenz plastidärer TPTs (Weber et al, 2006), über die Lokalisation und Substratspezifität der putativen Translokatoren ist jedoch bislang nahezu nichts bekannt.

1.5 *Phaeodactylum tricornutum* als Modellorganismus

Diatomeen sind ökologisch höchst relevante Mikroalgen, die für ca. 20% der globalen Photosyntheseleistung verantwortlich sind (Armbrust, 2009). Sie machen einen Großteil des marinen Phytoplanktons aus, kommen jedoch auch in limnischen Ökosystemen vor. Ein besonderes Charakteristikum ist ihre aus Silikat bestehende Zellwand (Vardi et al, 2008), weshalb Diatomeen auch als Kieselalgen bezeichnet werden. Taxonomisch sind Diatomeen den Heterokontophyten zuzuordnen, welche sich, wie bereits unter 1.2.2 dargestellt, durch den Besitz einer von vier Hüllmembranen umgebenen komplexen Plastide rhodophytischen Ursprungs auszeichnen.

Phaeodactylum tricornutum gilt als prominenter Modellorganismus für molekularbiologische Studien an Heterokontophyten. Das Genom der Diatomee ist vollständig sequenziert und weist eine Größe von ca. 27,4 Megabasen auf (Bowler et al, 2008). Ein weiterer Vorteil ist, dass die Diatomee genetisch stabil transformiert werden kann (Apt et al, 1996; Zaslavskaja et al, 2000), was *P. tricornutum* zu einem geeigneten Organismus für Protein-Lokalisationsstudien macht. Vor allem der Proteinimport in die komplexe Plastide kann in *P. tricornutum* in geeigneter Weise studiert werden, da sich die Subkompartimente nach Expression von z.B. GFP (grün fluoreszierendes Protein)-Fusionsproteinen fluoreszenzmikroskopisch problemlos unterscheiden lassen (Abb. 1.6).

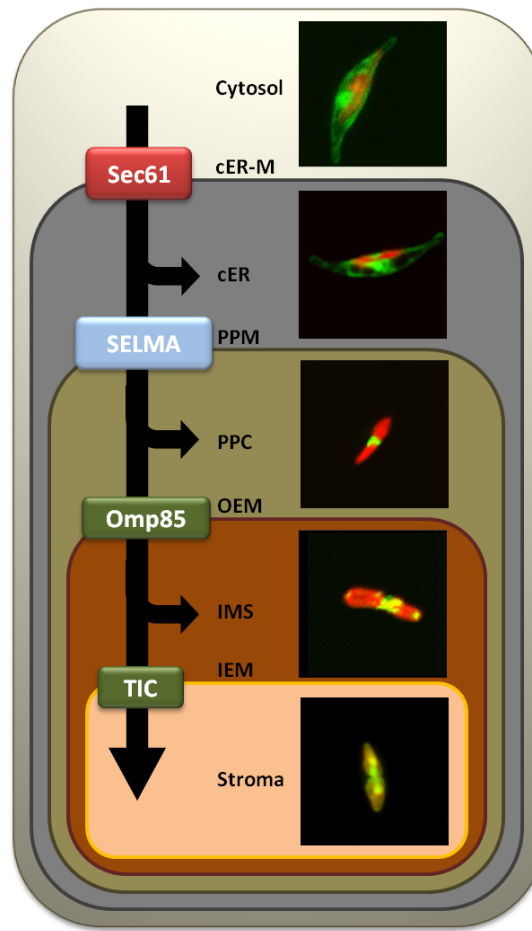


Abb. 1.6 Importstudien von GFP-Fusionsproteinen in der komplexen Plastide von *P. tricornutum*. Die Diatomee *P. tricornutum* stellt einen geeigneten Organismus für Proteinimportstudien in komplexe Plastiden dar, da die Subkompartimente fluoreszenzmikroskopisch unterschieden werden können. cER(-M), *chloroplast ER(-Membran)*; PPM, *periplastidäre Membran*; PPC, *periplastidäres Kompartiment*; OEM, *plastid outer envelope membrane*; IMS, *intermembrane space*; IEM, *plastid inner envelope membrane*; rot, *Plastidenautofluoreszenz*; grün, *eGFP-Fluoreszenz*. Abbildung modifiziert nach Bullmann et al. (2010).

2. Zielsetzung

Das PPC (*periplastidal compartment*, periplastidäres Kompartiment) der Diatomee *Phaeodactylum tricornutum* stellt ein auf natürliche Weise minimiertes eukaryotes Cytoplasma dar, das auf das ehemalige Cytosol des rhodophytischen Endosymbionten zurückgeht. Über die Funktion und den Proteingehalt dieses Kompartiments ist bisher, aufgrund von fehlenden biochemischen Methoden zur Isolierung des PPCs, nur wenig bekannt. Im Rahmen dieser Arbeit soll die Proteinkomposition des PPC von *P. tricornutum* analysiert werden, um Kenntnisse über grundlegende Prozesse hinsichtlich Transport, Metabolismus, Regulation und struktureller Organisation in diesem Kompartiment zu erlangen. Die Analyse soll auf Basis eines kombinierten *in silico/in vivo*-Ansatz erfolgen, indem potentielle PPC-Proteine zunächst anhand N-terminaler Zielsteuerungssequenzen bioinformatisch identifiziert und im Weiteren über die Expression von GFP-Fusionsproteinen *in vivo* hinsichtlich ihrer tatsächlichen Lokalisation untersucht werden.

In einem weiteren Aspekt dieser Arbeit soll die metabolische Kommunikation von Wirt und Symbiont untersucht werden. In diesem Kontext soll im speziellen die Präsenz von Triosephosphat/Phosphat-Translokatoren in den Membranen der komplexen Plastide von *P. tricornutum* aufgeklärt werden, welche eine mögliche Verbindung zwischen metabolischen Prozessen des Symbionten und des Wirts gewährleisten könnten. Identifizierte Kandidaten sollen dazu über die Generierung von GFP-Fusionsproteinen hinsichtlich ihrer subzellulären Lokalisation und Orientierung in der jeweiligen Zielmembran analysiert werden.

3. Ergebnisse

3.1. Charakterisierung des periplastidären Kompartiments von *Phaeodactylum tricornutum*

Die bioinformatische Analyse der Proteinkomposition des periplastidären Kompartiments (PPC) von *P. tricornutum* erfolgte anhand zweier unterschiedlicher Ansätze (siehe 6.2.5). In einem „gerichteten“ Ansatz wurde durch Datenbankabgleiche (Blast-Suche) gezielt nach typischen Faktoren für ein eukaryotes Cytoplasma gesucht, die mit einer PPC-spezifischen Zielsteuerungssequenz (*bipartite topogenic signal*, BTS) ausgestattet sind (6.2.5). Der „nicht gerichtete“ Ansatz zielte dahingegen auf die Detektion eines möglichst umfassenden Spektrums von löslichen PPC-Proteinen mit gezieltem Fokus auf ihre spezifischen N-terminalen Zielsteuerungssequenzen ab (6.2.5).

Insgesamt wurden bei dem „gerichteten Ansatz“ ca. 500 Genmodelle detektiert, überprüft und die kodierten Proteine hinsichtlich ihrer putativen Lokalisation, wie unter 6.2.5 beschrieben, eingeordnet (Tab. 8.1, Abb. 3.1). Aus diesem Datensatz wurden 48 Proteine, aufgrund der Detektion einer putativen PPC-spezifischen BTS (Tab. 8.2), über Präsequenz- bzw. FL (*full length*, volle Länge) -eGFP Fusionskonstrukte *in vivo* auf ihre subzelluläre Lokalisation hin untersucht. Für 27 davon konnte eine PPC-Lokalisation bestätigt werden, während insgesamt 21 der Fusionsproteine eine andere (nicht PPC) Lokalisation aufzeigten. Von diesen sogenannten „falsch-Positiven“ wurden 13 im ER bzw. dem sekretorischen Weg, sechs im Mitochondrium, sowie jeweils ein Protein im Stroma der Plastide und im Cytosol detektiert (siehe Tab. 8.1). Die Resultate dieses Ansatzes zeigten, dass bezüglich der Analyse mit den derzeit verfügbaren bioinformatischen Tools (u.a. TargetP), vor allem hinsichtlich der Vorhersage PPC-spezifischer Transitpeptid-ähnlicher Sequenzen von Organismen mit komplexen Plastiden, enorme Defizite bestehen. Für nur insgesamt ca. 55 % der 48 Proteinmodelle konnte eine bioinformatisch vorhergesagte PPC-Lokalisation auch durch die anschließende *in vivo* Lokalisation validiert werden.

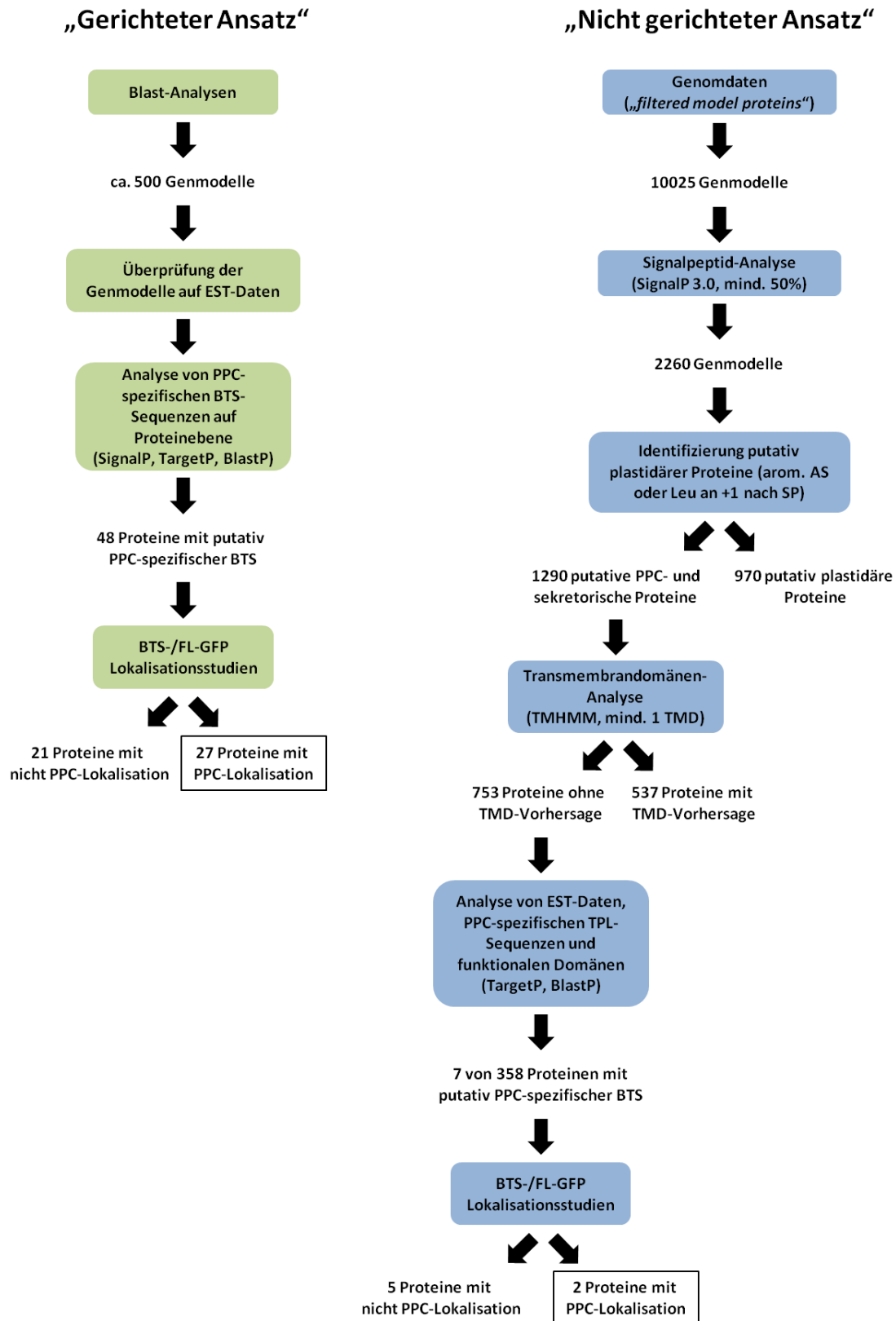


Abb. 3.1 Flussdiagramm zu den beiden kombinierten *in silico/in vivo* Ansätzen zur Identifizierung periplastidärer Proteine in *P. tricornutum*. SP, Signalpeptid; TPL, *transit peptide-like sequence*; BTS, *bipartite topogenic signal* (SP + TPL); FL, *full length*; TMD, Transmembrandomäne; EST, *expressed sequence tag*. Zur näheren Erläuterung siehe Text und 6.2.5.

Beim „nicht gerichteten“ Ansatz wurde damit begonnen alle vorhergesagten Proteinmodelle (10025) aus der *P. tricornutum* Genomdatenbank bioinformatisch, wie unter 6.2.5 beschrieben, auf die Präsenz einer PPC-spezifischen BTS hin zu überprüfen und geeignete Kandidaten im Anschluss *in vivo* zu lokalisieren. Durch diesen Ansatz konnten bioinformatisch 1290 Sequenzen mit vorhergesagtem SP ermittelt werden (putative PPC- und sekretorische Proteine), wobei 753 davon als lösliche Proteine ohne Transmembrandomäne (Vorhersage durch TMHMM) eingeordnet wurden. Davon wurden bis dato 358 Proteinmodelle, wie unter 6.2.5 beschrieben, näher untersucht. Da ein Großteil dieser Fraktion entweder aufgrund vorhergesagter funktioneller Domänen durch NCBI BlastP als ER-lokalisiert bzw. sekretorisch oder plastidär eingestuft oder bei einem beträchtlichen Anteil keine Homologie zu anderen Proteinsequenzen festgestellt werden konnte, wurden lediglich 7 dieser Sequenzen lokalisiert. Davon konnte nur für zwei Proteine eine PPC-Lokalisation durch *in vivo* Experimente bestätigt werden, während die restlichen 5 entweder im ER (2) oder dem Plastidenstroma (3) detektiert werden konnten (Tab. 8.2).

Abhängig vom Grad der Eingrenzbarkeit der putativen N-terminalen Zielsteuerungssequenzen (maures Protein der Präsequenz folgend konserviert/nicht konserviert), wurde in beiden Ansätzen entweder die Präsequenz (BTS) oder das Protein in voller Länge als GFP-Fusion exprimiert (Tab. 8.1). Insgesamt wurde in beiden Ansätzen die *in vivo* Lokalisation von 55 Proteinen mit putativ PPC-spezifischer BTS überprüft. Davon konnten jedoch nur 29 Proteine tatsächlich als PPC-lokalisiert bestätigt werden. Darunter fanden sich u.a. weitere putative SELMA-Faktoren, Chaperone, Proteasen und proteasomale 20S-Komponenten, sowie Proteine des Teilungsapparates der komplexen Plastide und Enzyme mit metabolischer Relevanz. Es gelang im Rahmen dieser Arbeit weder typische Vesikelfaktoren (COPI/II, Clathrin, ESCRT, SNAREs, Rabs, GEFs, GAPs, etc.) noch klassische Cytoskelettelemente (Actin, Tubulin, Intermediärfilamente, Myosin, Ankyrin) oder regulatorische bzw. Signaltransduktionskomponenten (Kinasen, Phosphatasen, Cryptochrome, Glutathion-System) in einer PPC-lokalisierten Version in *P. tricornutum* zu detektieren, während Wirtsfaktoren zumeist präsent waren (Tab. 8.1). Die Ergebnisse beider Ansätze sind in den folgenden Abschnitten dieses Kapitels dargestellt und sind nach funktioneller Einordnung der identifizierten Proteine gegliedert.

3.1.1 Proteinimport - SELMA-Faktoren

SELMA gilt als ERAD-abgeleitete Translokationsmaschinerie für nucleuskodierte Proteine über die zweite, periplastidäre Membran (PPM) der komplexen Plastide von *P. tricornutum* (Hempel et al, 2009; Sommer et al, 2007). Um die bisherigen Kenntnisse über das SELMA-System in *P. tricornutum* zu re-evaluieren und ggf. neue Faktoren zu identifizieren, wurde, wie unter 6.2.5 beschrieben, nach weiteren im Genom kodierten Komponenten gesucht. Dabei wurden alle bereits identifizierten SELMA- und ERAD-Komponenten der Wirtszelle aus *P. tricornutum*, sowie weitere bekannte ERAD-Faktoren aus anderen Organismen (*S. cerevisiae*, *C. merolae*, *A. thaliana*) gegen die *P. tricornutum* Genomdatenbank abgeglichen.

Die Suche ergab, dass neben der symbiontspezifischen AAA-ATPase sCdc48 (im Weiteren als Cdc48-1 bezeichnet) ein weiteres putativ symbiontspezifisches Cdc48 Protein (sCdc48-2) im PPC von *P. tricornutum* existiert (Abb. 3.2A und 3.3A). sCdc48-2 weist, verglichen mit sCdc48-1 und dem wirtsspezifischen hCdc48, eine ähnliche Domänenstruktur auf (siehe Domänenalignment Abb. 3.2A). Vor allem die beiden AAA-Domänen, sowie die N-terminale CDC48_N-Domäne sind hoch konserviert (siehe 8.3.1). Die bei hCdc48 durch die NCBI *Conserved Domain Search* (NCBI-CDS) vorhergesagte CDC48_2-Domäne, ein Teilbereich der N-terminalen Substraterkennungsdomäne von Cdc48 (Coles et al, 1999), konnte bei den beiden symbiontischen sCdc48-Proteinen nicht detektiert werden.

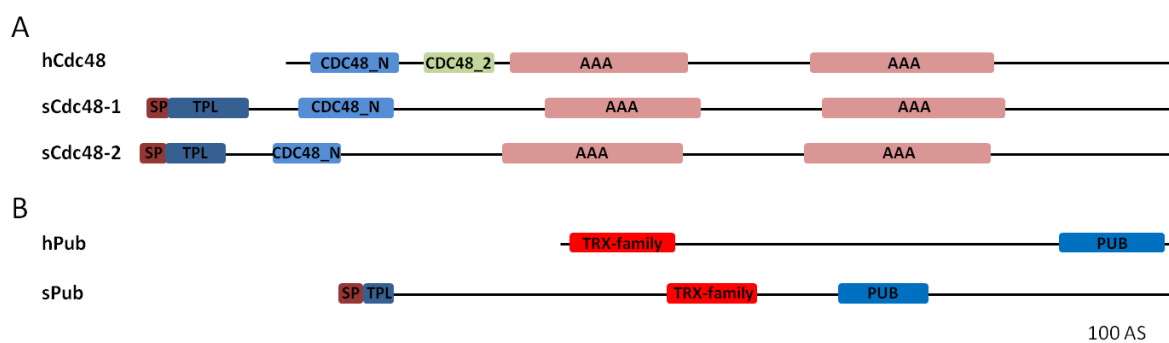


Abb. 3.2 Domänenalignments der Proteinprimärsequenzen von wirts- und symbiontspezifischen Cdc48- und Pub-Proteinen aus *P. tricornutum*. Struktureller Vergleich der durch NCBI *Conserved Domain Search* detektierten Domänen der wirts- und symbiontspezifischen Cdc48- (A) und Pub-Proteine (B). SP, Signalpeptid; TPL, Transitpeptid-ähnliche Sequenz (*transit peptide-like sequence*); CDC, *cell division cycle*; AAA, *ATPase associated with various cellular activities*; TRX, *Thioredoxin*; PUB, *peptide:N-glycanase/UBA- or UBX-containing proteins*.

Die Größte Varianz der Proteinsequenzen finden sich, abgesehen von den nicht-konservierten, putativen BTS-Sequenzen der symbiontischen Proteine, am N- und C-Terminus, den Bereichen die für die Bindung substratrekrutierender (N) und substratprozessierender (C) Cofaktoren bekannt sind (Stolz et al, 2011). Ähnlich dem symbiontspezifischen sCdc48-1, findet sich bei sCdc48-2 eine putativ PPC-spezifische BTS mit der Schnittstelle NHG-SVG (SignalPv3.0 HMM) am Übergang von Signalpeptid (SP) zu TPL (Abb. 3.2A und Tab. 8.2). *In vivo* Lokalisationsstudien mit der ermittelten BTS-Sequenz, sowie der AS-Sequenz von sCdc48-2 in voller Länge (FL, *full length*) fusioniert mit eGFP, ergaben eine für das PPC typische „Blob-ähnliche“ Lokalisationsstruktur („*blob-like structure*“, BLS) (Gruber et al, 2007; Kilian & Kroth, 2005) für beide Fusionskonstrukte (Abb. 3.3A).

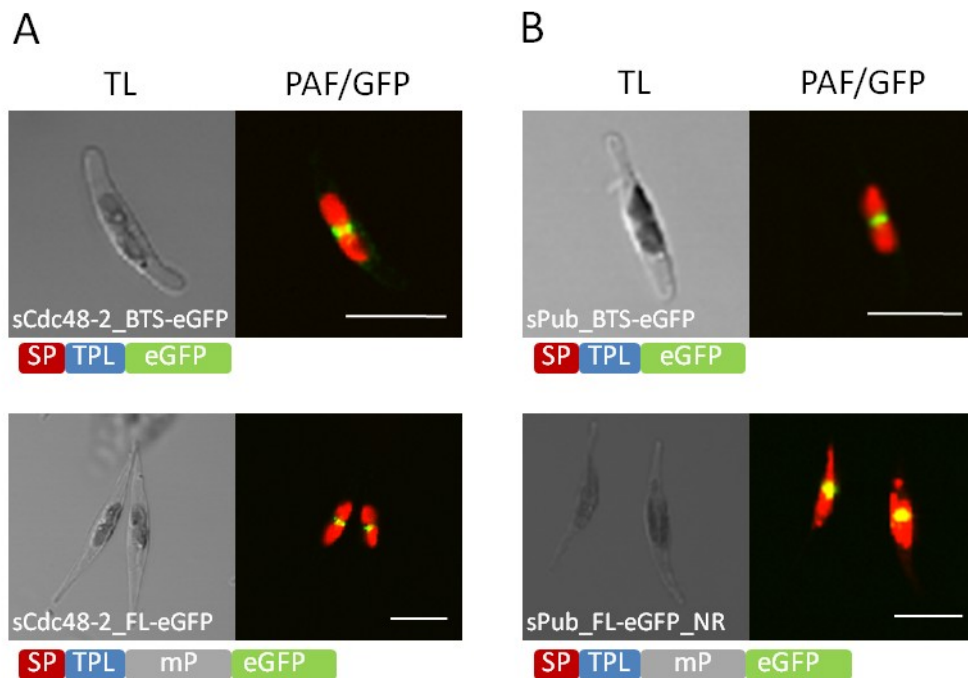


Abb. 3.3 *In vivo* Lokalisation von sCdc48-2 und sPub im PPC von *P. tricornutum*. Die Expression der BTS- und volle Länge Sequenzen (FL, *full length*) von sCdc48-2 (A) und sPub (B) fusioniert an eGFP resultierte in einer PPC-charakteristischen Lokalisationsstruktur („*blob-like structure*“, BLS). Mit Ausnahme von sPub-FL-eGFP (pPha_NR) wurden alle Fusionskonstrukte aus dem pPha-T1 Vektor exprimiert. TL, *transmitted light* (Durchlicht); PAF, Plastidenautofluoreszenz; GFP, *green fluorescent protein*. SP, Signalpeptid; TPL, *transit peptide-like sequence*; mP, matures Protein. Maßstab = 10 μ m.

Ein weiterer putativer SELMA-Faktor im PPC von *P. tricornutum* konnte mit sPub identifiziert werden. Dieses Protein weist, neben einer putativ PPC-spezifischen N-terminalen BTS (Schnittstelle: VDA-GLH; Tab. 8.2), eine Thioredoxin-, sowie eine PUB-Domäne (*peptide:N-glycanase/UBA- or UBX-containing proteins*) auf (Abb. 3.2B). Im Gegensatz zu

den zuvor genannten sCdc48-Proteinen findet sich bei sPub kein direktes wirtsspezifisches Pendant im Cytosol von *P. tricornutum*. Im Zuge der bioinformatischen Analyse konnte lediglich ein Protein mit gleicher Domänenkomposition (hPub, Abb. 3.2B), jedoch ohne signifikante AS-Sequenzidentität identifiziert werden. Um die Lokalisation von sPub in *P. tricornutum* zu überprüfen, wurden jeweils die BTS, sowie die FL-Sequenz des Proteins als eGFP-Fusionskonstrukt *in vivo* exprimiert. Wie Abbildung 3.3B verdeutlicht, konnte bei beiden Fusionsproteinen unter dem KLSM (konfokales Laser Scanning Mikroskop) eine Lokalisation im PPC beobachtet werden.

3.1.1.1 Interaktionsstudien mit sCdc48-1/2 und sPub mittels Split-GFP System

Nachdem für sCdc48-2 und sPub eine PPC-Lokalisation nachgewiesen werden konnte (3.1.1), wurden im Weiteren Interaktionstest mit den beiden Proteinen und auch sCdc48-1 durchgeführt um Hinweise auf ihre potentiellen Funktionen zu gewinnen. Dazu wurde das Split-GFP System genutzt (6.2.3.10.3), welches bereits in vorangehenden Studien in *P. tricornutum* erfolgreich angewendet werden konnte (Hempel et al, 2009).

Zunächst wurde über Split-GFP eine potentielle Interaktion von sCdc48-1 mit sCdc48-2 getestet. Es ist bekannt, dass sich Cdc48-Proteine zur Ausübung ihrer Funktionen für gewöhnlich in homohexameren Ringstrukturen organisieren (Wang et al, 2003). Vorangehende Analysen einer homotypischen Interaktion des PPC-lokalisierten sCdc48-1 über Split-GFP (duale Expression von sCdc48-1_FL_N-GFP und sCdc48-1_FL_C-GFP, Nitrat-induziert über pPha_Dual_2xNR) konnten dies jedoch nicht belegen (Franziska Hempel, persönliche Kommunikation). Es wurde daher überprüft, ob sCdc48-1 und sCdc48-2 möglicherweise befähigt sind Heterooligomere auszubilden oder in anderer Weise miteinander interagieren. In diesem Kontext wurde N-GFP als C-terminale Fusion von sCdc48-1 und C-GFP C-terminal an sCdc48-2 fusioniert und in *P. tricornutum* exprimiert (6.2.3.10.3). Abbildung 3.4A verdeutlicht, dass dies zu GFP-Fluoreszenz im PPC von *P. tricornutum* führte, was auf eine Interaktion beider Proteine hindeutet.

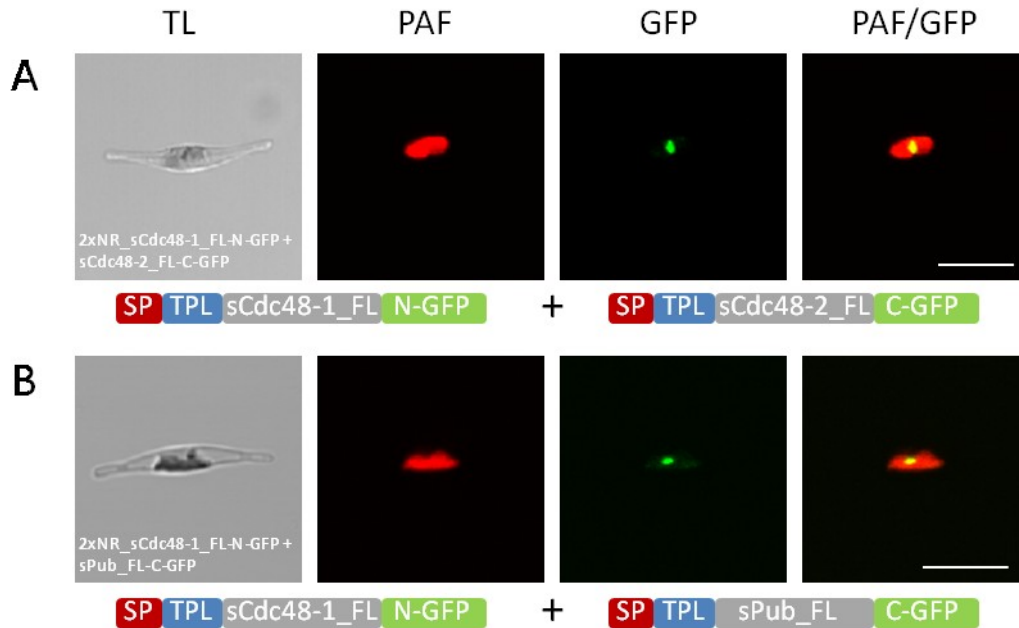


Abb. 3.4 *In vivo* Interaktionsstudien von sCdc48-1/2 und sPub im PPC von *P. tricornutum* über das Split-GFP System. (A) Die Expression von sCdc48-1_FL-N-GFP und sCdc48-2_FL-C-GFP über pPha_Dual_2xNR führte zu einer GFP-Fluoreszenz im PPC. (B) Nach Expression von sCdc48-1_FL-N-GFP zusammen mit sPub_FL-C-GFP über pPha_Dual_2xNR konnte eine PPC-spezifische Fluoreszenz von GFP beobachtet werden. TL, *transmitted light* (Durchlicht); PAF, *Plastidenautofluoreszenz*; GFP, *green fluorescent protein*. Maßstab = 10 µm.

Da PUB-Domänen als Cdc48-Interaktionsmodule bekannt sind (Allen et al, 2006; Madsen et al, 2009; Zhao et al, 2007), wurde im Weiteren eine mögliche Interaktion von sPub mit sCdc48-1 bzw. sCdc48-2 im PPC von *P. tricornutum* analysiert. Dazu wurde sPub sowohl C-terminal fusioniert mit C-GFP zusammen mit sCdc48-1_FL-N-GFP, als auch C-terminal fusioniert mit N-GFP zusammen mit sCdc48-2_FL-C-GFP in *P. tricornutum* exprimiert. Wie Abbildung 3.4B veranschaulicht, resultierte die Expression von sPub_FL-C-GFP und sCdc48-1_FL-N-GFP in GFP-Fluoreszenz, was eine Interaktion beider Proteine im PPC bestätigt. Eine Interaktion von sCdc48-2 und sPub konnte dahingegen bisher experimentell nicht nachgewiesen werden.

3.1.2 Molekulare Chaperone

Damit Proteine nach erfolgter Translation im Cytosol ihre korrekte native Konformation einnehmen können, werden zumeist molekulare Chaperone benötigt, die aktiv am Faltungsprozess mitwirken. Außerdem scheint die Erhaltung der Transportkompetenz

(Faltungsverhinderung) vor bzw. während der Translokation von Präproteinen durch Chaperone insbesondere im PPC für die Passage der OEM eine wichtige Rolle zu spielen (Bullmann et al, 2010). Um die Präsenz weiterer Chaperone (neben sHsp70) im PPC von *P. tricornutum* zu untersuchen, wurde die Genomdatenbank insbesondere auf kodierte Chaperonine und Proteine mit TPR-, DnaJ-, PDI- oder Peptidyl-Prolyl cis-trans Isomerase-Domäne, welche u.a. bei Hsp70-Cochaperonen (TPR, DnaJ) und anderen an der Faltung von Proteinen beteiligten Enzymen vorkommen, geprüft. Das Screening resultierte in der Detektion von insgesamt 61 putativen Chaperonen (Tab 8.1) von denen bei drei Kandidaten potentiell PPC-spezifische Zielsteuerungssequenzen ermittelt werden konnten (Tab. 8.2, Abb. 3.5).

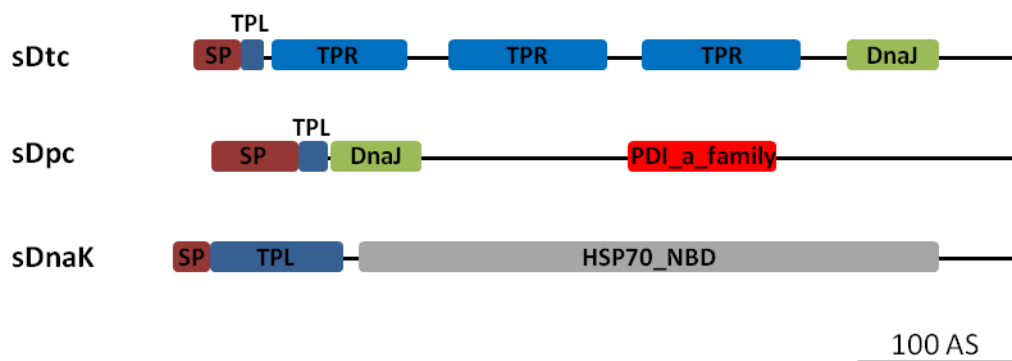


Abb. 3.5 Detektierte Domänen putativer PPC-Chaperone aus *P. tricornutum*. SP, Signalpeptid; TPL, Transitpeptid-ähnliche Sequenz; TPR, *tetratricopeptide repeats*; DnaJ, Hsp70 Interaktionsdomäne, typisch für Hsp70-Cochaperone; PDI_a_family, *protein disulfide isomerase* (PDIa) family; Hsp70_NBD, Hsp70 nucleotide binding domain; die Detektion putativer Domänen erfolgte mit NCBI-CDS.

Für zwei dieser Proteine (detektiert durch Blast-Analysen mit Scj1p aus *S. cerevisiae*) konnte durch NCBI-CDS eine DnaJ-Domäne detektiert werden, welche auf Proteinebene dafür bekannt ist die ATPase-Domäne von Hsp70-Proteinen zu regulieren und somit stabilisierend auf die Bindung von Hsp70 zu Substraten wirkt (Cyr et al, 1994; Laufen et al, 1999; Qiu et al, 2006). Eines dieser Proteine weist, neben der C-terminal gelegenen DnaJ-Domäne, drei TPR- (*tetratricopeptide repeats*) Motive auf (Abb. 3.5) und wurde daher, sowie aufgrund fehlender bekannter Homologe, als sDtc (*DnaJ and TPR domain-containing protein*) bezeichnet. Bei dem zweiten DnaJ-Protein liegt die DnaJ-Domäne N-terminal unmittelbar hinter der putativen BTS. Blast-Analysen ergaben, dass dieses Protein zusätzlich eine PDI- (Protein-Disulfid-Isomerase) Domäne im Mittelteil besitzt (sDpc, *DnaJ and PDI domain-containing protein*). Neben den beiden potentiell PPC-lokalisierten DnaJ-Proteinen konnten durch

bioinformatische Analysen eine Reihe weiterer Proteine mit DnaJ-Domäne identifiziert werden, die jedoch aufgrund bioinformatischer Vorhersagen als ER- oder mitochondriell-lokalisiert bzw. cytosolisch eingestuft wurden (Tab. 8.1).

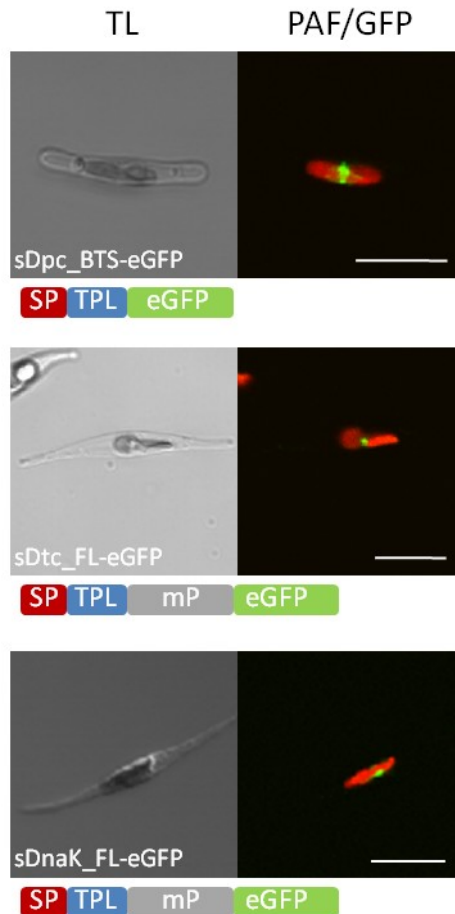


Abb. 3.6 *In vivo* Lokalisation neuer Chaperone im PPC von *P. tricornutum*. Die Expression der BTS von sDpc, sowie der FL-Sequenzen von sDtc (Moog, 2009) und sDnaK fusioniert mit eGFP resultierten in einer PPC-charakteristischen GFP-Fluoreszenz. TL, *transmitted light* (Durchlicht); PAF, Plastidenautofluoreszenz; GFP, *green fluorescent protein*; SP, Signalpeptid; TPL, *transit peptide-like sequence*; mP, matures Protein. Maßstab = 10 μ m.

Ein weiteres Chaperon-ähnliches Protein konnte über den „nicht gerichteten“ Ansatz identifiziert werden (6.2.5). Hierbei handelt es sich um ein Protein des prokaryoten Hsp70-Typs, DnaK. sDnaK weist eine durch NCBI-CDS vorhergesagte, für Hsp70-Proteine typische Nukleotidbindedomäne (NBD) auf. Eine typische Substratbindedomäne wurde bei sDnaK nicht erkannt. Lokalisationsstudien mit der BTS von sDpc, sowie den Sequenzen von sDtc (Moog, 2009) und sDnaK in voller Länge (FL) fusioniert an eGFP ergaben eine PPC-Lokalisation für alle drei Fusionsproteine (Abb. 3.6).

3.1.3 Proteasomale Komponenten und Proteasen

Eng assoziiert an das ERAD-System eukaryoter Zellen (vgl. SELMA) ist das 26S-Proteasom. Dieser Multikatalytische Proteinkomplex setzt sich aus zwei Hauptbestandteilen zusammen: dem 20S Proteasom Kernkomplex, bestehend aus 28 Untereinheiten (α und β), die sich in heteroheptameren Ringen (α 1-7, β 1-7) in einer Fassstruktur anordnen (Baumeister et al, 1998), sowie dem regulatorischen 19S Partikel (mindestens 18 Untereinheiten, u.a. AAA-ATPasen) (Gallastegui & Groll, 2010; Murata et al, 2009). Das 26S-Proteasom dient u.a. der Proteolyse bzw. dem Abbau defekter polyubiquitinerter Proteine im Zuge der ER-assoziierten Degradation (ERAD). Darüberhinaus besitzt dieses eine Funktion bei der Regulation zellulärer Prozesse durch gezielte Degradation spezifischer Proteine im Kontext räumlicher und zeitlicher Kontrollmechanismen (Voges et al, 1999).

Die Sequenzierung des Nucleomorphgenoms der Cryptophyte *Guillardia theta* ergab, dass dieser für 21 Untereinheiten eines 26S Proteasomkomplexes kodiert, was somit auf das Vorkommen eines eukaryoten Proteasoms im PPC von *G. theta* hindeutete (Douglas et al, 2001). Um die mögliche Präsenz eines Proteasoms im PPC von *P. tricornutum* zu untersuchen, wurde die *P. tricornutum* Genomdatenbank mit bekannten AS-Sequenzen des *S. cerevisiae* 26S-Proteasomkomplexes abgeglichen. Die Suche ergab, neben der Detektion eines nahezu vollständigen Sets an proteasomalen (19S und 20S) Wirtskomponenten (siehe Tab. 8.1), dass einige 20S-Faktoren in doppelter Ausführung im Genom von *P. tricornutum* kodiert sind. Dabei handelt es sich um fünf alpha (α 2, α 3-1, α 3-2, α 7-1 und α 7-2) und vier beta Untereinheiten (β 2, β 3, β 6 und β 7), für welche jeweils eine putativ PPC-spezifische BTS detektiert werden konnte. Die Identifizierung von α 3-1 gelang dabei über den nicht gerichteten Ansatz, während alle anderen 20S-Untereinheiten über Blast-Analysen (siehe 6.2.5) detektiert wurden.

Auffällig bei den potentiell PPC-lokalisierten 20S-Untereinheiten waren ihre putativen N-terminalen Zielsteuerungssequenzen (BTS). Zum einen wurde für α 7-2, sowie für β 6 anstelle eines SP durch den HMM-Algorithmus von SignalP 3.0 eine Signalankersequenz (SA) erkannt, während die Detektion der Signalpeptide der restlichen 20S-Faktoren eindeutig war. Zum anderen sind die als TPL eingestuftten Bereiche der meisten 20S-Komponenten mit einer Länge von 4 bis 16 AS sehr kurz, d.h. der konservierte Bereich der Proteine beginnt

nahezu direkt hinter der vorhergesagten SP-Schnittstelle. Ausnahmen bilden dabei $\alpha 7$ -2 und $\beta 7$, welche putative TPL Sequenzen von 39 bzw. 96 AS aufweisen (siehe 8.2).

Die Klassifizierung der jeweiligen Untereinheiten nach alpha 1-7 und beta 1-7 erfolgte mittels NCB-CDS. Für alle 20S-Faktoren konnte dabei eine charakteristische *Ntn* (*N-terminal nucleophile*)-*hydrolase superfamily*-Domäne detektiert werden (nicht gezeigt). Um die vorhergesagte PPC-Lokalisation der potentiell symbiontspezifischen proteasomalen 20S-Komponenten zu überprüfen, wurden bisher sechs der neun Proteine entweder als BTS- oder FL-eGFP Fusionskonstrukte in *P. tricornutum* exprimiert. Abbildung 3.7 verdeutlicht, dass für alle sechs Proteine ($\alpha 3$ -1, $\alpha 7$ -1, $\alpha 7$ -2, $\beta 2$, $\beta 7$) bzw. ihre Präsequenzen ($\beta 6$) fusioniert mit eGFP eine PPC-spezifische Fluoreszenzstruktur (BLS) beobachtet werden konnte.

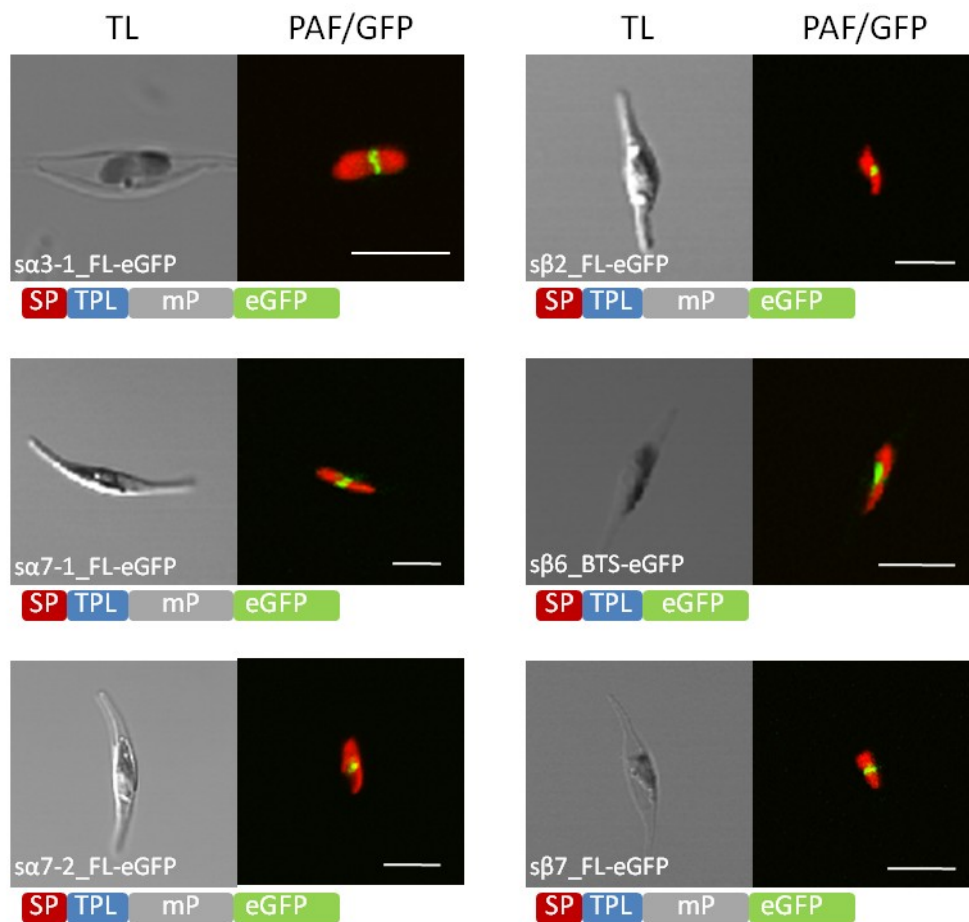


Abb. 3.7 In vivo Lokalisation proteasomaler 20S-Faktoren im PPC von *P. tricornutum*. Die Expression von BTS- ($\beta 6$) und FL-Sequenzen ($\alpha 3$ -1, $\alpha 7$ -1, $\alpha 7$ -2, $\beta 2$, $\beta 7$) putativ symbiontspezifischer proteasomaler 20S-Komponenten fusioniert an eGFP ergab für alle Fusionsproteine eine PPC-charakteristische Lokalisationsstruktur. TL, *transmitted light* (Durchlicht); PAF, *Plastidenautofluoreszenz*; GFP, *green fluorescent protein*; SP, *Signalpeptid*; TPL, *transit peptide-like sequence*; mP, *matures Protein*. Maßstab = 10 μ m.

Es ist besonders zu erwähnen, dass das bioinformatische Screening keinerlei Hinweise auf die Präsenz von Komponenten eines regulatorischen 19S-Partikels im PPC von *P. tricornutum* ergab, welcher beim eukaryoten 26S-Proteasom für die Erkennung ubiquitiniierter Substrate, sowie deren Entfaltung und Weiterleitung in den 20S-Kernkomplex zuständig ist (Voges et al, 1999). Auf Seiten des Wirts konnte im Gegensatz dazu ein nahezu vollständiges Set an Proteinen des 19S regulatorischen Partikels identifiziert werden (siehe Tab. 8.1).

Da im PPC von *P. tricornutum* nur ein unvollständiges Proteasom detektiert werden konnte und eine proteolytische Funktion somit fraglich ist, wurde zusätzlich nach weiteren Proteasen gesucht, die eine Funktion bei der Prozessierung bzw. Umsetzung von Proteinen einnehmen könnten. Dabei gelang es über die direkte Suche nach Annotation von Genmodellen bzw. spezifischen Dömanen (*P. tricornutum* Genomdatenbank Phatr2, *SEARCH*, „Trypsin“) und zusätzlichen Abgleichen der so identifizierten AS-Sequenzen gegen die *P. tricornutum* Genomdatenbank (BlastP), sieben Trypsin-ähnliche Serinproteasen (Tlp, *trypsin-like serine protease*) zu detektieren. Alle Proteine besitzen eine durch die NCBI-CDS erkannte und konservierte *trypsin-like serine protease*-Domäne (Tryp_SPc, Abb. 3.8). Für sTlp3 konnte zusätzlich eine PT-Domäne detektiert werden, welche eine Abfolge von XPTX-Tetrapeptiden (X = beliebige AS) mit unbekannter Funktion darstellt. Außerdem handelt es sich bei sTlp2 und sTlp3 wahrscheinlich um Membranproteine, da bei beiden eine C-terminale Transmembrandomäne (TMD) durch TMHMM vorhergesagt wurde. Bei sechs der sieben Tlp-Proteine wurde eine putative PPC-spezifische BTS detektiert (Abb. 3.8, Tab. 8.2), wohingegen eines der sieben Proteine keine N-terminale Zielsteuerungssequenz aufzuweisen scheint. Die Proteine wurden dementsprechend als hTlp (wirtsspezifisch) bzw. sTlp1 bis sTlp6 (symbiontspezifisch) benannt.

Wie in Abbildung 3.9 dargestellt, besitzen alle Tlp-Proteine aus *P. tricornutum* die an der katalytischen Triade von Serin-Proteasen beteiligten konservierten Aminosäuren (His57, Asp102, Ser195). Abweichungen werden insbesondere bei den Substratspezifitäts-Determinanten deutlich. Demnach findet sich bei hTlp, sTlp2, sTlp3, sTlp4 und sTlp5 ein für Trypsin-ähnliche Substratspezifität charakteristisches Aspartat an Position 189, während sTlp1 und sTlp6 an dieser Position Glycin aufweisen.

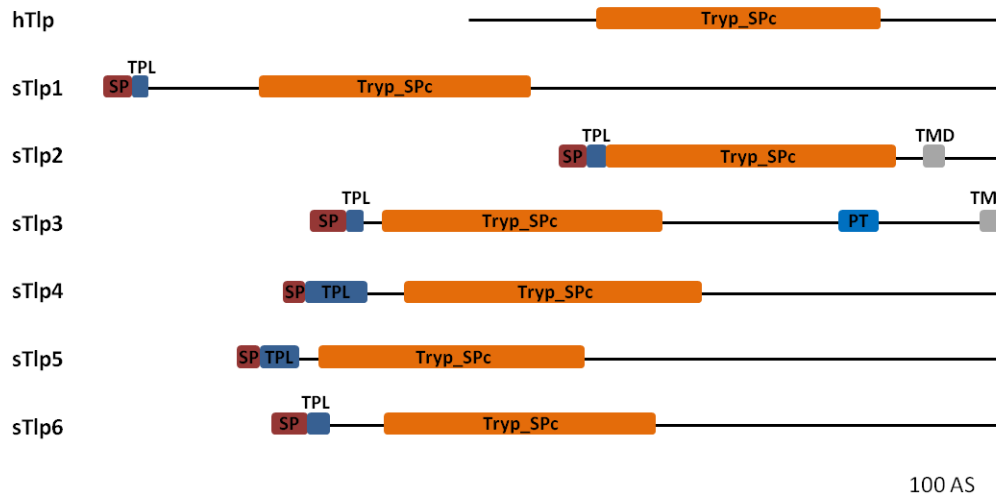


Abb. 3.8 Domänenkomposition der Tlp-Proteine aus *P. tricornutum*. Alle Proteine besitzen eine durch NCBI-CDS erkannte Tryp_SpC-Domäne (*trypsin-like serine protease domain*). Für sechs der sieben Tlp-Proteine konnte eine putative, PPC-spezifische BTS detektiert werden (sTlp1-6), wohingegen hTlp keine N-terminale Zielsteuerungssequenz aufweist und somit als wirtsspezifisches Protein eingeordnet wurde. SP, Signalpeptid; TPL, Transitpeptid-ähnliche Sequenz; Tryp_SpC, *trypsin-like serine protease domain*; PT, Proteindomäne unbekannter Funktion, bestehend aus einer Abfolge von X-Prolin-Threonin-X Tetrapeptiden (X = beliebige Aminosäure); TMD, Transmembrandomäne (Vorhersage durch TMHMM).

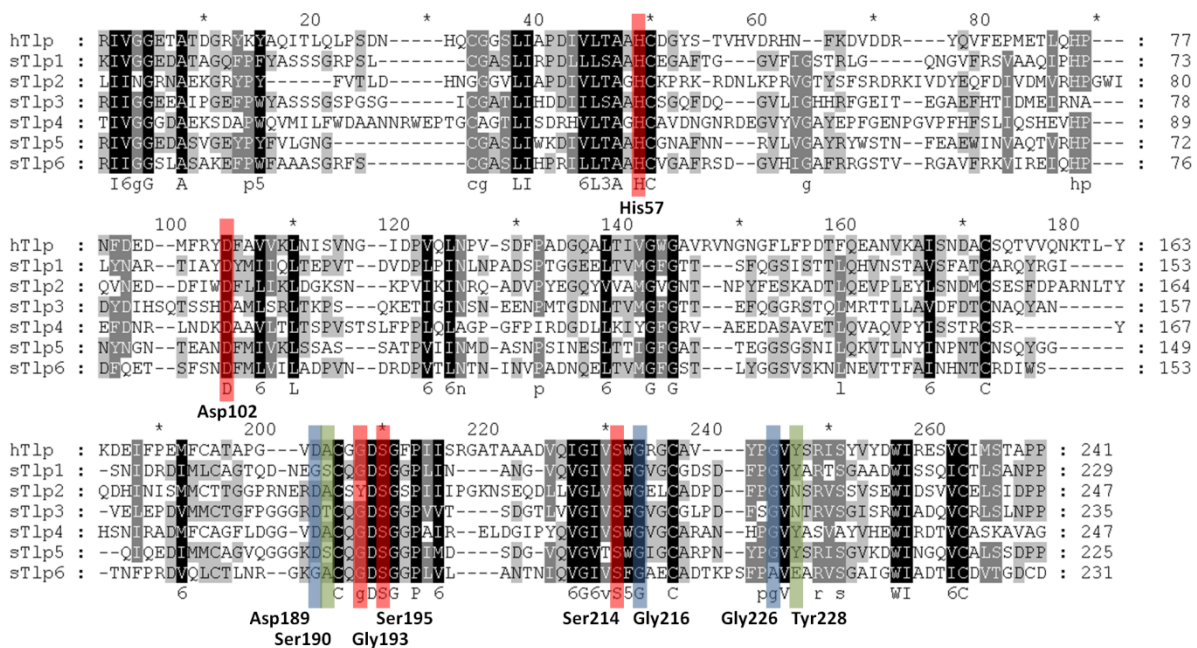


Abb. 3.9 Primärsequenz-Alignment der aktiven Zentren und stark konservierten Sequenzbereiche der Tlp-Proteine aus *P. tricornutum*. Die an der katalytischen Triade beteiligten AS His57, Asp102 und Ser195 sind bei allen Tlp-Proteinen konserviert. Unterschiede werden bei den Substratspezifität-Determinanten (blau bzw. grün) deutlich und offenbaren bei hTlp, sTlp2, sTlp3, sTlp4 und sTlp5 ein für Trypsin typisches Aspartat an Position 189, wobei diese Position bei sTlp1 und sTlp6 jeweils von Glycin eingenommen wird. Rot, am Katalysemechanismus beteiligte Reste; blau, Position primärer Determinanten für Substratspezifität; grün, Position sekundärer Determinanten für Substratspezifität. Die markierten AS-Positionen beziehen sich auf die aus Trypsin bekannte Nummerierung nach Perona und Craig (1995).

Um die durch bioinformatische Analysen vorhergesagte PPC-Lokalisation der sTlp-Proteine in *P. tricornutum* zu überprüfen, wurden *in vivo* Lokalisationsstudien der FL- bzw. BTS-Sequenzen von bisher fünf der sechs Proteine fusioniert an eGFP durchgeführt. Wie Abbildung 3.10 verdeutlicht konnte für alle Fusionskonstrukte eine PPC-spezifische Lokalisation beobachtet werden.

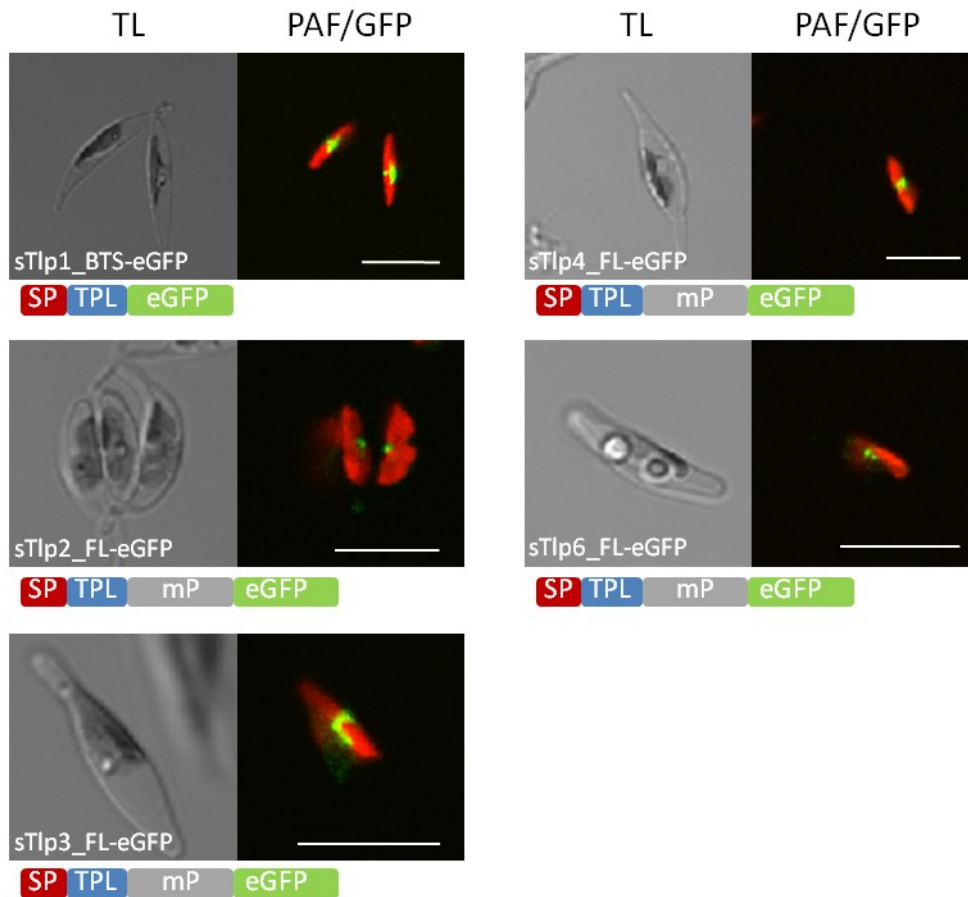


Abb. 3.10 *In vivo* Lokalisation symbiontischer Trypsin-ähnlicher Serin-Proteasen (sTlp) im PPC von *P. tricornutum*. Die Expression von BTS- (sTlp1) und FL-Sequenzen (sTlp2, sTlp3, sTlp4, sTlp6) der sTlp-Proteine fusioniert an eGFP ergab für alle Fusionsproteine ein PPC-spezifisches Fluoreszenzmuster. TL, *transmitted light* (Durchlicht); PAF, *Plastidenautofluoreszenz*; GFP, *green fluorescent protein*; SP, *Signalpeptid*; TPL, *transit peptide-like sequence*; mP, *matures Protein*. Maßstab = 10 μ m.

Zusätzlich zu den zuvor beschriebenen Proteinen wurden in dem Screening nach Proteasen drei ClpS- und vier FtsH-Proteine identifiziert, die jedoch nach bioinformatischen Analysen nicht im PPC, sondern dem Cytosol, dem Plastidenstroma oder den Mitochondrien lokalisieren (Tab. 8.1).

3.1.4 Plastidenteilungsfaktoren

Die Teilung komplexer Plastiden, welche von vier Membranen umgeben werden, erfolgt in einem hoch regulierten Prozess, der bisher nur in Ansätzen verstanden ist. In primären Plastiden ist dieser Prozess besser untersucht. Dort sind mehrere ringförmige Strukturen an der konzentrischen Teilung der von zwei Hüllmembranen umgebenen Organellen beteiligt. Diese sogenannten Plastidenteilungs- (*plastid-dividing*, PD) Ringe formieren sich in primären Plastiden an der cytosolischen Seite der OEM, der stromalen Seite der IEM und zumindest in Rotalgen (*C. merolae*) auch im Intermembranraum und wurden in elektronenmikroskopischen Studien als elektronendichte Strukturen beschrieben (McAndrew et al, 2001; Miyagishima et al, 2001). PD-Ringstrukturen konnten auch in komplexen Plastiden beobachtet werden (*P. tricornutum*), dort allerdings bisher nur im ehemaligen Cytosol des Endosymbionten (PPC) um die OEM herum, jedoch nicht im inneren der Plastide (Hashimoto, 2005). Erst kürzlich gelang es die Zusammensetzung des cytosolischen PD-Rings in *C. merolae* aufzuklären. Demnach besteht dieser aus Polyglucanfilamenten (Glucose), die im Cytosol von einer Glycogenin-ähnlichen Glycosyltransferase der GT8-Familie, genannt PDR1 (*plastid-dividing ring 1*), um die OEM konstruiert werden (Yoshida et al, 2010).

Zusätzlich zu den PD-Ringen wirken an der Teilung primärer Plastiden eine Reihe weiterer Proteinfaktoren mit, die sich z.T. ebenfalls in ringartigen Strukturen formieren. Zentrale Komponenten bildet dabei die GTPasen DRP5B (*dynamain-related protein*) im Cytosol und FtsZ (*filamenting temperature-sensitive mutant Z*) in der Plastide. Des Weiteren sind in Grünalgen und höheren Pflanzen eine Reihe akzessorischer bzw. regulatorischer Faktoren, wie z.B. Komponenten des bakteriellen Min-Systems, beteiligt (Miyagishima & Kabeya, 2010; Yang et al, 2008).

3.1.4.1 Bioinformatische Analysen und *in vivo* Lokalisationen putativer Komponenten des Teilungsapparats der komplexen Plastide

Um die Präsenz möglicher Komponenten einer Teilungsmaschinerie der komplexen Plastide im PPC von *P. tricornutum* zu untersuchen, wurden bioinformatische Analysen unternommen. Vorwiegend wurde dabei nach den aus der Rotalge *C. merolae* bekannten,

cytosolischen Faktoren Dnm2 (DRP5B) (Miyagishima et al, 2003) und PDR1 (Yoshida et al, 2010), jedoch auch nach stromalen Komponenten, wie FtsZ, Min-Proteinen und anderen an der Teilung primärer Plastiden beteiligten Faktoren (Yang et al, 2008), gesucht. Mittels BlastP Analysen gelang es dabei ein Dynamin-ähnliches Protein, im Folgenden als sDrp bezeichnet, in *P. tricornutum* zu identifizieren, welches eine hohe Sequenzidentität (42%) mit dem an der Teilung der primären Plastide beteiligten Dnm2 aus *C. merolae* aufweist (siehe Alignment 8.3.1). sDrp besitzt die für DRP5B-Proteine typische DLP_1-Domäne (Abb. 3.11A), welche auch in Cm_Dnm2 vorhanden ist und die für die GTPase-Funktion essentiellen AS-Reste beinhaltet (Marks et al, 2001). Am N-Terminus von sDrp konnte eine putative PPC-spezifische BTS mit einer Länge von 124 AS detektiert werden (Tab. 8.2). Diese, sowie die FL-Sequenz von sDrp, wurden im Weiteren durch *in vivo* Lokalisationsstudien auf ihre subzelluläre Lokalisation hin untersucht. Abbildung 3.12A verdeutlicht, dass für beide Fusionsproteine eine PPC-spezifische GFP-Lokalisation beobachtet werden konnte.

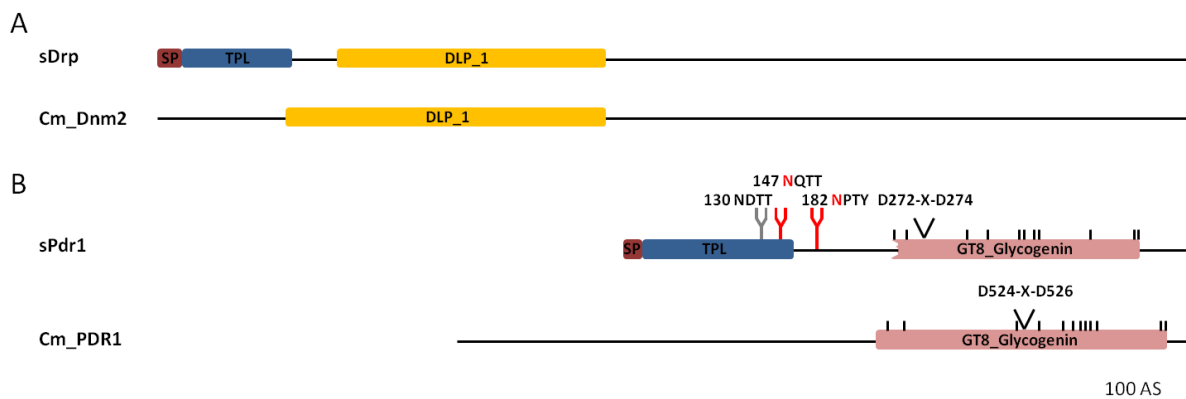


Abb. 3.11 Domänenalignement putativer Komponenten des Teilungsapparates der komplexen Plastide von *P. tricornutum*. (A) Gegenüberstellung von sDrp aus *P. tricornutum* und dem an der Teilung der primären Plastide beteiligten Dnm2 aus der Rotalge *Cyanidioschyzon merolae*. (B) Domänenvergleich von sPdr1 aus *P. tricornutum* mit dem PD-Ring aufbauenden PDR1 aus *C. merolae*. Die GT8_Glycogenin-Domäne von sPdr1 wird durch NCBI-CDS als N-terminal unvollständig erkannt. Durch NetNGlyc1.0 vorhergesagte N-Glycosylierungsstellen sind in Rot (hohe Wahrscheinlichkeit) bzw. Grau (geringe Wahrscheinlichkeit) dargestellt. Das bei Glycogeninen konservierte DXD-Motiv ist durch ein offenes Pfeil markiert, während die schwarzen Balken Tyrosinreste in der Glycogenin-Domäne andeuten, die für eine potentielle O-Glycosylierung in Frage kommen. SP, Signalpeptid; TPL, *transit peptide-like sequence*; DLP_1, *dynamamin-like protein*; GT8, *glycosyltransferase family eight*.

Hinsichtlich der bioinformatischen Suche nach Homologen zu Cm_PDR1, der PD-Ring konstruierenden, cytosolischen Glycosyltransferase aus *C. merolae* (Yoshida et al, 2010), in *P. tricornutum*, konnten über BlastP zunächst keine positiven Ergebnisse erzielt werden. Jedoch gelang es über die direkte Suche nach Genmodell-Annotation (*P. tricornutum*

Genomdatenbank Phatr2, *SEARCH*, „Glycogenin“) fünf für Glycogenin-ähnliche Proteine kodierende Gene in der Genomdatenbank zu identifizieren (Tab. 8.1). Für eines dieser fünf Proteine, fortan als sPdr1 bezeichnet, konnte eine potentielle, PPC-spezifische BTS (ca. 158 AS) identifiziert werden (8.2, Abb. 3.11B), wobei die Länge der putativen TPL aufgrund fehlender Konservierung am N-Terminus nicht eindeutig bestimmt werden konnte. Für die restlichen vier wurden durch SignalP entweder ein SP (GT8-1), Signalanker (GT8-4, GT8-5) oder keine N-terminalen Zielsteuerungssequenz (GT8-3) vorhergesagt. Während die Homologie von sPdr1 zu Cm_PDR1 auf Sequenzebene nur gering ausfällt (siehe Alignment 8.3.3), besitzen beide Proteine jedoch eine homologe Domänenstruktur (Abb. 3.11B). Die durch NCBI-CDS detektierte GT8_Glycogenin-Domäne befindet sich jeweils am C-Terminus der Proteine, sPdr1 ist jedoch gegenüber Cm_PDR1 N-terminal verkürzt (Abb. 3.11B). Beide Proteine verfügen über ein bei Glycogeninen konserviertes DXD-Motiv innerhalb der GT8_Glycogenin-Domäne (Lomako et al, 2004).

Glycogenine sind als selbst-glucosylierende Enzyme bekannt, welche in der Lage sind, Glucosereste über eine, einen Tyrosinrest involvierende, O-glycosidische Bindung kovalent an sich zu binden (Lomako et al, 2004). Mögliche Tyrosine, die für eine O-Glycosylierung von sPdr1 und Cm_PDR1 in Frage kommen sind in Abbildung 3.11B dargestellt. Da sPdr1, im Gegensatz zu dem cytosolischen, putativen Pendant aus *C. merolae*, das ER passieren muss, wurde die Primärsequenz zusätzlich bioinformatisch auf potentielle N-Glycosylierungsstellen untersucht. Eine Analyse mit NetNGlyc1.0 ergab, dass für sPdr1 solche an Position 130, 147 und 182 vorhergesagt werden konnten (Abb. 3.11B). Während Position 130 den von NetNGlyc1.0 vorgegebenen Schwellenwert von 0,5 nicht überschreitet (N-Glycosylierung unwahrscheinlich), ist die Vorhersage für eine N-Glycosylierung an Position 147 und 182 mit Werten von 0,6 bzw. 0,59 positiv. Im Falle von N182 befindet sich jedoch an Position 183 ein Prolin, was an dieser spezifischen Position vorkommend dafür bekannt ist N-Glycosylierung zu verhindern (Bause, 1983). Somit ist eine potentielle N-Glycosylierung von sPdr1 an Position 147 am wahrscheinlichsten, diese befindet sich jedoch noch innerhalb des TPL-Bereichs der putativen BTS.

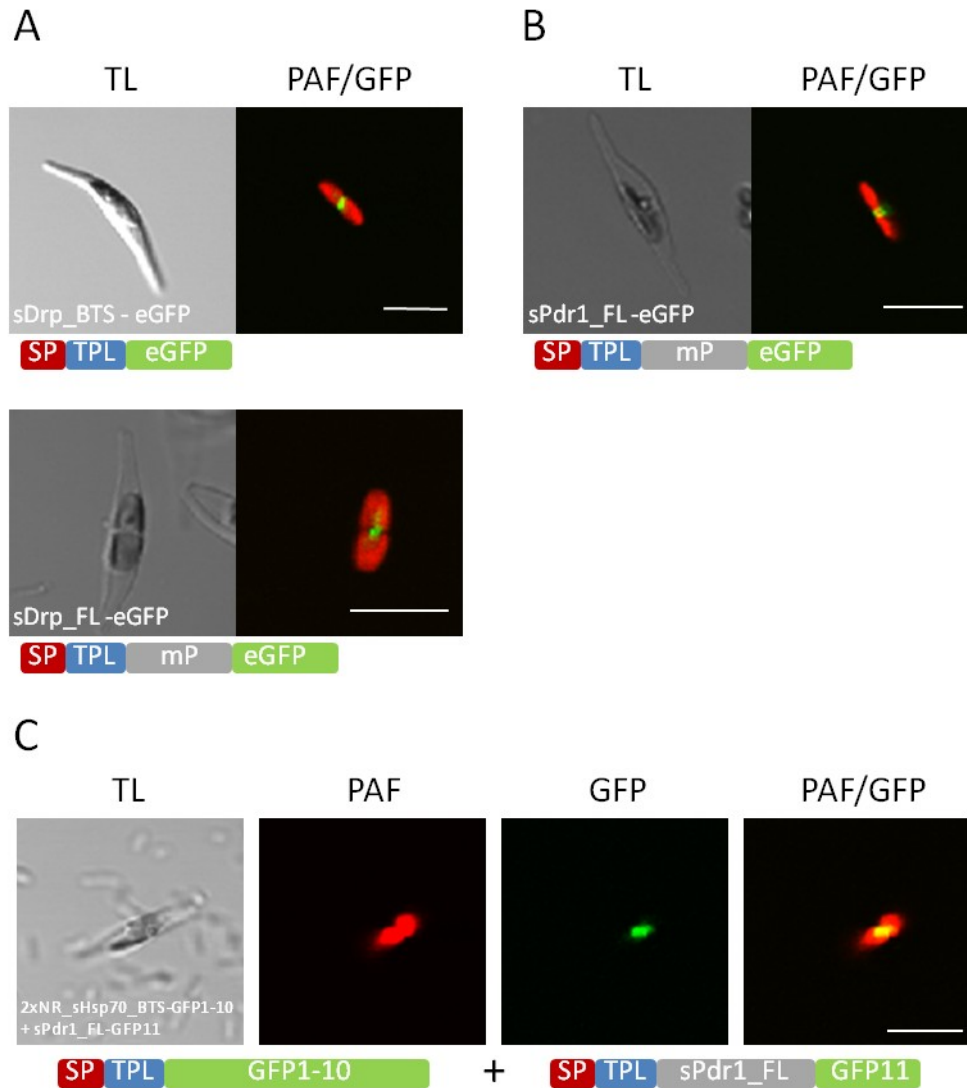


Abb. 3.12 *In vivo* Lokalisation putativer Komponenten des Teilungsapparates der komplexen Plastide im PPC von *P. tricornutum*. (A) Die Expression der BTS-, sowie FL-Sequenz von sDrp fusioniert an eGFP resultierte jeweils in einer PPC-Lokalisation. (B) Nach Expression eines Fusionskonstrukts bestehend aus der FL-Sequenz von sPdr1 und eGFP konnte eine GFP-Fluoreszenz im Bereich des PPC der komplexen Plastide beobachtet werden. (C) Die Nitrat-induzierte Expression (6 h) der BTS von sHsp70 fusioniert mit GFP1-10 und sPdr1_FL fusioniert mit GFP11 (Self-Assembling-GFP System) resultiert in einer GFP-Fluoreszenz im PPC. TL, *transmitted light* (Durchlicht); PAF, Plastidenautofluoreszenz; GFP, *green fluorescent protein*; SP, Signalpeptid; TPL, *transit peptide-like sequence*; mP, matures Protein. Maßstab = 10 μ m.

Um schließlich die vorhergesagte Lokalisation von sPdr1 in *P. tricornutum* zu überprüfen, wurden *in vivo* Lokalisationsstudien mit der FL-Sequenz von sPdr1 fusioniert an eGFP durchgeführt. Abbildung 3.12B verdeutlicht, dass eine PPC-Lokalisation für das exprimierte Fusionskonstrukt beobachtet werden konnte, welche sich in manchen Zellen in einer ringförmigen Struktur um die Engstelle der primären Plastide darstellte. Die PPC-Lokalisation von sPdr1 wurde außerdem mit dem Self-Assembling-GFP System (SA-GFP) überprüft

(6.2.3.10.2). Die bereits in anderen Studien in diesem Kontext erfolgreiche genutzte BTS von sHsp70 (PPC-Marker) (Bullmann et al, 2010; Hempel et al, 2009) wurde dazu verwendet GFP_S1-10 in das PPC von *P. tricornutum* zu dirigieren (sHsp70_BTS-GFP1-10), wobei sPdr1 in voller Länge als Fusionsprotein vor GFP_S11 exprimiert wurde (sPdr1_FL-GFP11). Nach sechsstündiger Induktion durch Zugabe von Nitrat (Hempel et al, 2009) wurden die Klone unter dem KLSM betrachtet. Wie Abbildung 3.12C veranschaulicht, konnte GFP-Fluoreszenz im Bereich des PPC beobachtet werden, was eine PPC-Lokalisation von sPdr1 erneut bekräftigt.

Hinsichtlich der Identifizierung weiterer Faktoren der Teilungsmaschinerie der komplexen Plastide von *P. tricornutum* ließen sich bisher bioinformatisch nur drei für vermeintlich stromale FtsZ-Proteine kodierende Gene detektieren (8.1). Gene kodierend für weitere Komponenten, wie Proteine des plastidären Min-Systems, sowie ARC3, ARC6, PARC6, MCD1, PDV1, PDV2 und GC1 scheinen zu fehlen, bzw. konnten weder im Kerngenom noch im Plastidengenom der Diatomee identifiziert werden.

3.1.4.2 Analyse des Glycosylierungsstatus von sPdr1

Nach erfolgreicher Lokalisation von sPdr1 im PPC von *P. tricornutum*, wurde anschließend untersucht, ob sPdr1, ähnlich wie sein putatives Pendant PDR1 aus *C. merolae*, zur Autoglucosylierung befähigt ist (Yoshida et al, 2010). Positive *P. tricornutum* Klone, die mit dem kodierenden Gen für sPdr1_FL-eGFP in dem Nitrat-induzierbaren pPha_NR-Vektor transformiert worden waren und eine PPC-Lokalisation des exprimierten Fusionsproteins unter dem KLSM zeigten, wurden dazu zunächst einem Expressionstest unterzogen. Von sieben solcher Klone wurde jeweils Gesamtproteinextrakt hergestellt (6.2.3.1) und mittels Antikörper gegen GFP wurde anschließend der Klon mit der stärksten sPdr1_FL-eGFP-Expression über SDS-PAGE und Western Blot ermittelt (siehe Abb. 8.4). Dieser Klon wurde darauffolgend zur Immunopräzipitation von sPdr1_FL-eGFP verwendet (6.2.3.6). Nach erfolgter Aufreinigung des Fusionsproteins wurde die Hälfte des Eluats mit PNGase F behandelt (6.2.3.9), während die andere Hälfte unbehandelt blieb. Dieses Verfahren diente dazu potentielle N-Glycosylierungen von der vermuteten O-Glycosylierung am Glycogenin-ähnlichen sPdr1 unterscheiden zu können.

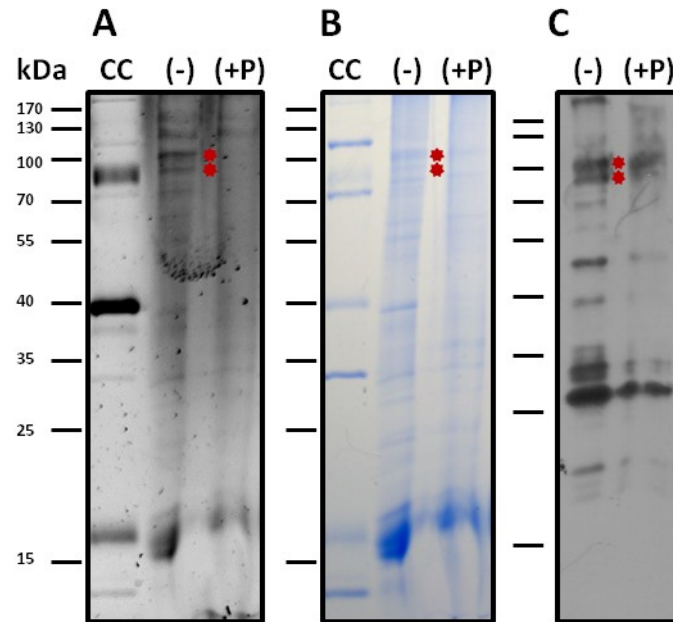


Abb. 3.13 Analyse des Glycosylierungsstatus des PPC-lokalisierten sPdr1_FL-eGFP aus *P. tricornutum*. (A) *Glyco-Stain*. Nach Anfärbung von potentiellen Glycosylresten und Visualisierung unter UV-Licht konnten in Relation zu der Kontrolle (CC Marker) zwei deutliche Banden in der unbehandelten Fraktion (-) von sPdr1_FL-eGFP beobachtet werden (Markierung). In der PNGase-behandelten Fraktion (+P) konnten diese Signale nicht detektiert werden. (B) Coomassie-Färbung. Das zum *Glyco-Stain* eingesetzte SDS-Gel wurde in einem Folgeschritt mit Coomassie gefärbt um alle im Gel aufgetrennten Proteinbanden zu visualisieren. Die Markierung deutet auf die korrespondierenden Banden aus dem *Glyco-Stain* hin. (C) Western-Blot. Ein Teil der PNGase-behandelten (+P) und unbehandelten (-) Fraktionen (jeweils 40 %) wurde zum Western Blot eingesetzt. Die Detektion von sPdr1_FL-eGFP erfolgte mit einem Antikörper gegen GFP (6.2.3.5 und 6.1.6). Die Markierung zeigt die entsprechenden Banden aus A und B. Das über PeptideMass ermittelte Molekulargewicht von sPdr1_FL-eGFP beträgt 88,6 kDa, für die prozessierte Form (ohne BTS) wurde anhand der unter 3.1.4.1 und 8.2 beschriebenen Kalkulation ein Molekulargewicht von ca. 71 kDa ermittelt. kDa, Kilodalton; CC, CandyCane Marker (Kontrolle), bestehend aus glycosylierten und unglycosylierten Proteinen (siehe 6.1.9).

Im Weiteren wurde das immunpräzipitierte Fusionsprotein (PNGase F behandelt und unbehandelt) auf einem SDS-Gel aufgetrennt und einem *Glyco-Staining* unterzogen (6.2.3.8). Dieses höchst sensitive Verfahren beruht auf der spezifischen Färbung Perjodat-oxidierter Zuckerreste an Glycoproteinen mittels UV-visuallisierbarem Fluoreszenzfarbstoffs. Wie Abbildung 3.13A verdeutlicht, konnte nach dem *Glyco-Staining* ein Signal für die unbehandelte Fraktion des immunpräzipitierten sPdr1_FL-eGFP detektiert werden. Korrespondierende Banden, die in Relation mit den Molekulargewichten des Markers unprozessiertem (88,5 kDa) bzw. prozessiertem (ca. 71 kDa) Fusionsprotein entsprechen (Banden erscheinen höher als kalkuliert), konnten auch bei einer anschließenden Coomassie-Färbung und dem Western Blot beobachtet werden (Abb. 3.13B/C). Wurde das rekombinante Fusionsprotein jedoch mit PNGase F behandelt, blieb das nach dem *Glyco-Stain* erwartete

Signal aus (Abb. 3.13A). Bei Betrachtung der Coomassie-Färbung (Abb. 3.13B) und des Western-Blots (Abb. 3.13C) wird zudem deutlich, dass die Behandlung mit PNGase F zu einer Reduzierung der Proteinkonzentration führte. Außerdem zeigten sich in beiden Fraktionen Degradationsprodukte des aufgereinigten sPdr1_FL-eGFP Fusionsproteins, für die keine Signale im *Glyco-Stain* detektiert werden konnten.

3.1.5 Lipidtransfer, Fettsäurebiosynthese und Glycero(phospho)lipidmetabolismus

Die äußere Membrane primärer Plastiden (OEM) beinhaltet neben den typischen Glycolipiden auch Phospholipide, wie z.B. Phosphatidylcholin (Dorne et al, 1985). Es wird vermutet, dass der sich während der primären Endosymbiose ereignende Verlust des früheren cyanobakteriellen Lipopolysaccharids und Lipoproteins Voraussetzung für den Import von Proteinen von Seiten des Wirts war und somit einen wichtigen Schritt zur Etablierung des Cyanobakteriums als Organell (primäre Plastide) darstellte (Cavalier-Smith, 2000). Der Transport von Phospholipiden vom Ort ihrer Synthese am ER der Wirtszelle zur äußeren Lipidschicht der OEM primärer Plastiden könnte dabei über Lipidtransferproteine erfolgt sein (Cavalier-Smith, 2000). Auf dieser Annahme basierend, wurde daher bioinformatisch nach Lipidtransfer-Proteinen (LTPs) gesucht, die möglicherweise eine Versorgung der ehemals primären Plastide mit ER-synthetisierten Phospholipiden gewährleisten könnten.

Dabei gelang es fünf Homologe zu dem aus *Saccharomyces cerevisiae* bekannten cytosolischen Lipidtransferprotein Sec14p in *P. tricornutum* zu identifizieren, wovon für eines (sSec14) eine potentielle PPC-spezifische BTS mit einer Länge von ca. 100 AS (Tab. 8.2) detektiert werden konnte. Bei den restlichen vier Homologen (Tab. 8.1) handelt es sich nach bioinformatischen Analysen um cytosolische Proteine. Sec14-Proteine sind dafür bekannt Phosphatidylinositol sowie Phosphatidylcholin zu binden, die Fähigkeit des aktiven Lipidtransfers konnte für Sec14 jedoch bisher nur *in vitro* nachgewiesen werden (Griac, 2007; Saito et al, 2007). Bei sSec14 aus *P. tricornutum* konnte durch NCBI-CDS eine Sec14p-charakteristische CRAL_TRIO-Domäne detektiert werden (nicht gezeigt), die Sequenzidentität zu Sec14p beträgt 27 %. Lokalisationsstudien, bei denen die FL-Sequenz von sSec14 als eGFP-Fusionsprotein in *P. tricornutum* exprimiert wurde, resultierten in einer PPC-spezifischen Fluoreszenzstruktur (Abb. 3.14) (Moog, 2009).

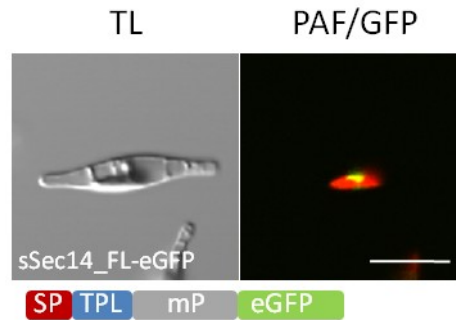


Abb. 3.14 *In vivo* Lokalisation von sSec14 im PPC von *P. tricornutum*. Die Expression des sSec14_FL-eGFP Fusionsproteins in *P. tricornutum* führte zu einer PPC-Lokalisation (Moog, 2009). TL, *transmitted light* (Durchlicht); PAF, Plastidenautofluoreszenz; GFP, *green fluorescent protein*; SP, Signalpeptid; TPL, *transit peptide-like sequence*; mP, matures Protein. Maßstab = 10 μ m.

Um des Weiteren die Präsenz möglicher Faktoren der Fettsäure- bzw. Glycero(phospho)lipidbiosynthese im PPC der Diatomee zu analysieren, die mit sSec14 in Verbindung stehen könnten, wurde ein ergänzendes bioinformatisches Screening unternommen. Wie Tabelle 8.1 verdeutlicht, konnten dabei jedoch keine Enzyme mit putativ PPC-spezifischen Zielsteuerungssequenzen detektiert werden. Die identifizierten Proteine wurden durch die angewandten Vorhersagealgorithmen in Kombination mit BlastP-Analysen entweder als cytosolisch, mitochondriell, plastidär oder ER-lokalisiert klassifiziert (Tab. 8.1).

3.1.6 Metabolismus - Carboanhydrasen

Die Untersuchung der enzymatischen Verteilung des Kohlenhydratstoffwechsels in *P. tricornutum* war bereits Gegenstand vorheriger Studien (Gruber et al, 2009; Kroth et al, 2008). In einer der Studien konnte gezeigt werden, dass mit s6PGDH ein Enzym des oxidativen Pentosephosphatwegs (OPP) im PPC der Diatomee lokalisiert ist und das PPC somit in metabolische Prozesse involviert sein könnte (Gruber et al, 2009). Hinsichtlich der Präsenz weiterer metabolischer Enzyme im PPC liegen bisher allerdings keine experimentellen Hinweise vor. Um zu analysieren ob das PPC von *P. tricornutum* möglicherweise Carboanhydrasen (CA) im Kontext der Konzentrierung anorganischen Kohlenstoffs (CCM, *carbon concentrating mechanism*) beherbergt, wurde ein genomisches Screening nach CA-kodierenden Genen und ihren Produkten unternommen. Die Analyse ergab, dass im Genom der Diatomee, neben zwei bereits bekannten β -CAs (plastidäre

Lokalisation) und fünf α -CAs (unbekannte Lokalisation) (Kitao et al, 2008; Kroth et al, 2008; Tanaka et al, 2005), drei weitere CAs der γ -Familie kodiert sind (Tab. 8.1). Die Primärsequenzen der kodierten Proteine wurden im Folgenden mittels SignalP und TargetP auf die Präsenz N-terminaler Zielsteuerungssequenzen untersucht. Dabei konnten für alle α CAs (α Ca1-5) Signalpeptide vorhergesagt werden (Tab. 8.2), während bei den drei γ CAs putativ mitochondrielle Zielsteuerungssequenzen ermittelt werden konnten (TargetP). Die α CAs wurden daraufhin mittels TargetP und BlastP auf die Präsenz putativer TPL-Sequenzen bzw. nicht konservierter Bereiche, welche den ermittelten Signalpeptiden folgen könnten, überprüft. Es ergab sich, dass für vier der fünf α CAs aufgrund geringer Konservierung am N-Terminus (alle besitzen eine durch NCBI-CDS detektierte alpha_CA-Domäne im Mittelteil bzw. C-terminal) mögliche PPC-spezifische TPL-Sequenzen angenommen werden konnten, wohingegen eines der CA-Proteine (α Ca3) wegen der Präsenz der AS Leucin an der ersten Position der möglichen TPL als plastidär oder ER-lokalisiert eingestuft wurde (Tab. 8.1).

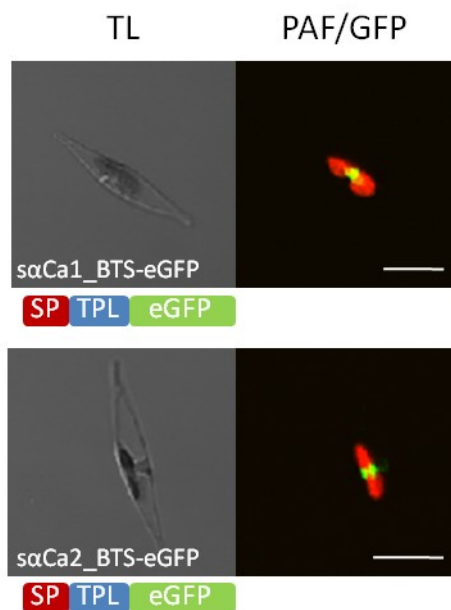


Abb. 3.15 *In vivo* Lokalisation putativer α -Carboanhydrasen im PPC von *P. tricornutum*. Nach Überexpression der putativen BTS-Sequenzen von $s\alpha$ Ca1 und $s\alpha$ Ca2 fusioniert an eGFP konnte eine PPC-charakteristische GFP-Fluoreszenz unter dem KLSM detektiert werden. TL, *transmitted light* (Durchlicht); PAF, Plastidenautofluoreszenz; GFP, *green fluorescent protein*; SP, Signalpeptid; TPL, *transit peptide-like sequence*. Maßstab = 10 μ m.

Mittels GFP-Lokalisationsstudien wurde daraufhin die Lokalisation der vier α CAs *in vivo* ermittelt. Dazu wurden die putativen Zielsteuerungssequenzen als eGFP-Fusion in *P. tricornutum* exprimiert (siehe Tab. 8.2). Wie Abbildung 3.15 verdeutlicht, konnte durch

diesen Ansatz für zwei α CAs (α Ca1 und α Ca2) eine PPC-Lokalisation ermittelt werden. Die Expression der anderen beiden α CA-eGFP-Fusionsproteine (α Ca1 und α Ca2) resultierte dahingegen in ER-typischen Fluoreszenzstrukturen (nicht gezeigt, Tab. 8.1).

3.1.7 Homologe zu Nucleomorph-kodierten Proteinen der Cryptophyte *Guillardia theta*

Cryptophyten besitzen gegenüber allen anderen bisher bekannten Organismengruppen mit komplexen Plastiden rhodophytischen Ursprungs einen genetisch aktiven Restzellkern im PPC, den sogenannten Nucleomorph (Zauner et al, 2000). In *Guillardia theta* beherbergt dieser drei Chromosomen mit einer Gesamtgröße von 551 kbp die für 464 putative Proteine kodieren. Der überwiegende Teil der Gene des Nucleomorphs kodiert für *housekeeping*-Faktoren (Transkription, Translation, etc.), sowie für ca. 30 plastidäre Proteine und ca. 219 putative Proteine mit bislang unbekannten Funktionen (Douglas et al, 2001).

Blast-Analysen ergaben, dass insgesamt 31 der der bisher unbekannten ORFs (*open reading frame*) des *G. theta* Nucleomorphs Homologien zu Proteinen aus *P. tricornutum* (20 ORFs) bzw. *C. merolae* (27 ORFs) aufweisen. Basierend auf dieser Grundlage wurde im Weiteren untersucht, ob Homologe zu den Nucleomorph-kodierten ORFs, die genomisch in *P. tricornutum* identifiziert werden konnten, im PPC der Diatomee lokalisieren. Für fünf der ORF-homologen Proteine aus *P. tricornutum* konnten über bioinformatische Analysen mögliche PPC-spezifische BTS-Sequenzen ermittelt werden (Tab. 8.2), während die restlichen Proteine zumeist keine N-terminalen Zielsteuerungssequenzen aufwiesen und anhand der Detektion putativer funktioneller Domänen (NCBI-CDS) dem Transkriptions- bzw. Translationsapparat der Wirtszelle zugeordnet werden konnten. Bei den identifizierten, putativ PPC-lokalisierten Proteinen handelt es sich um Homologe zu Gt_ORF532a (Nucleomorph, Chromosom I), Gt_ORF534 (Nm, Chr. I), Gt_ORF139 (Nm, Chr. II) und Gt_ORF261 (Nm, Chr. III) und Gt_ORF387 (Nm, Chr. III), welche im Weiteren als sORF532a, sORF534, sORF139, sORF261 und ORF387 bezeichnet werden.

Um die identifizierten Proteine näher zu charakterisieren bzw. um Informationen über mögliche Funktionen zu erhalten, wurden die Primärsequenzen einer Blast-Analyse unterzogen. Dabei stellte sich heraus, dass für die putativ symbiontischen ORFs aus *P. tricornutum* über NCBI-CDS bzw. BlastP keine konservierten funktionellen Domänen

detektiert werden konnten. Die BlastP-Analysen zeigten jedoch, dass, mit Ausnahme von sORF139, Homologe in Rhodophyta und Organismen mit komplexen Plastiden rhodophytischen Ursprungs existieren (siehe 3.2). Eine Überprüfung der ORF-Sequenzen aus *G. theta* mit NCBI-CDS ergab, dass für Gt_ORF532a eine polC-Domäne (Polymerase III, PolC) und für Gt_ORF139 eine DsbA- (oxidative Proteinfaltung) bzw. Thioredoxin-ähnliche Domäne detektiert werden konnte. Des Weiteren wurden auch die ORF-Homologen aus Rotalgen (*C. merolae*) auf die Präsenz funktioneller Domänen untersucht. Dabei zeigte sich, dass bei den Rotalgenproteinen mit Homologie zu ORF532a (CMQ054C), ORF534 (CMI197C) und ORF261 (CMD101C) eine RHD3-Domäne (*root hair defective*) identifiziert werden konnte. Diese putativen GTP-bindenden Domänen befinden sich hauptsächlich am N-Terminus und im Mittelteil der Rotalgenproteine, die Bereiche starker Sequenzkonservierung gegenüber den Homologen aus *P. tricornutum* (und *G. theta*) liegen jedoch vor allem C-terminal (siehe Alignments 8.3.5, 8.3.6 und 8.3.7).

Die identifizierte Domäne für das Rotalgenhomolog von Gt_ORF139 (CMR441C) war kongruent mit dem Protein aus *G. theta* (DsbA-Domäne), beide Proteine besitzen jedoch nur geringe Sequenzidentität mit sORF139 aus *P. tricornutum* (siehe Alignment 8.3.4), bei dem eine solche Domäne nicht detektiert werden konnte. Für ORF387 ließen sich auch bei den entsprechenden Proteinen aus *G. theta* (Gt_ORF387) und *C. merolae* (CMP057C) keine funktionellen Domänen identifizieren. Datenbankabgleiche mit der Primärsequenz von ORF387 deuteten darauf hin, dass homologe Proteine, außer in Organismen mit komplexen Plastiden die auf eine Rotalge zurückgehen, sowie Rotalgen selbst, hauptsächlich in *Vibrio spec.* (γ -Proteobakterien) vorkommen.

Um letztlich die Lokalisation der *P. tricornutum* Proteine mit Homologie zu den *G. theta* Nucleomorph-kodierten ORFs zu überprüfen, wurden *in vivo* Lokalisationsstudien durchgeführt. Nach Expression der FL-Sequenzen von sORF139, sORF261, sORF532a und sORF534 fusioniert mit eGFP konnte in allen Fällen eine „Blob-ähnliche“, PPC-typische GFP-Fluoreszenzstruktur beobachtet werden (Abb. 3.16). Im Gegensatz dazu ergab die Expression von ORF387_FL-eGFP in *P. tricornutum*, dass für das Fusionsprotein weder eine Lokalisation im PPC noch in anderen Kompartimenten der komplexen Plastide der Diatomee detektiert werden konnte (nicht gezeigt).

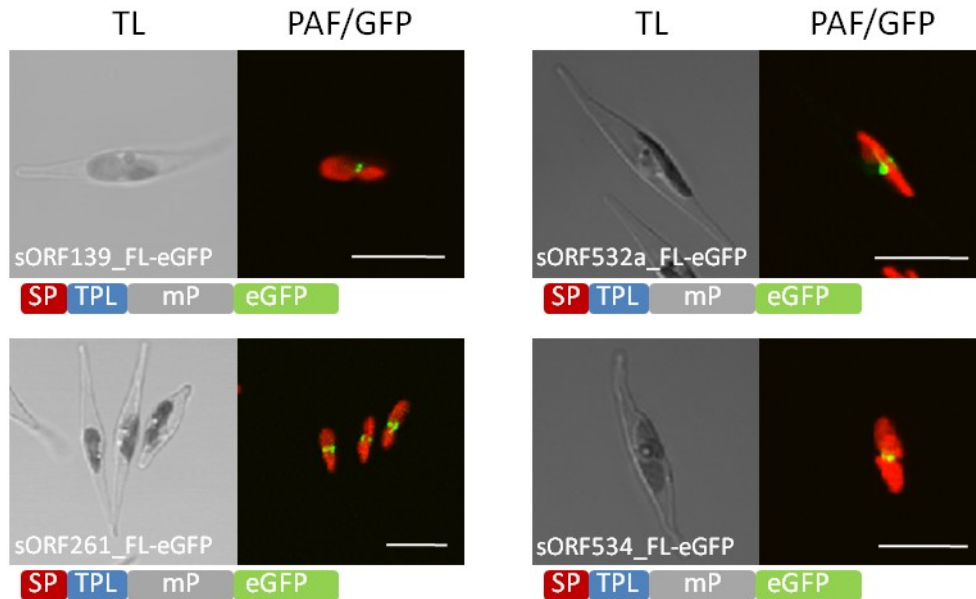


Abb. 3.16 *In vivo* Lokalisation homologer Proteine zu Nucleomorph-kodierten ORFs der Cryptophyte *Guillardia theta* im PPC von *P. tricornutum*. Die Expression der FL-Sequenzen von sORF139, sORF261, sORF532a und sORF534 aus *P. tricornutum* fusioniert mit eGFP resultierte in einer PPC-spezifischen Lokalisation der Fusionsproteine (Moog, 2009). TL, *transmitted light* (Durchlicht); PAF, Plastidenautofluoreszenz; GFP, *green fluorescent protein*; SP, Signalpeptid; TPL, *transit peptide-like sequence*; mP, matures Protein. Maßstab = 10 μ m.

3.1.8 Weitere Faktoren

Neben den oben beschriebenen PPC-lokalisierten Faktoren in *P. tricornutum*, konnten in den beiden *in silico* Ansätzen (6.2.5) weitere putative PPC-Komponenten identifiziert werden, bei denen eine Reihe diverser, potentieller Domänen vorhergesagt werden konnten (NCBI-CDS). Diese Faktoren ließen sich jedoch, zumeist aufgrund unbekannter Funktionen, bisher nicht in definierte Kategorien einordnen und wurden daher unter diesen gemeinschaftlichen Punkt zusammengefasst. Zu diesen Proteinen zählen u.a. Frustuline, Prolyl-4-Hydroxylasen, Pectinesterasen, sowie eine Reihe weiterer Faktoren (siehe Tab. 8.1, „Sonstige“).

Insgesamt konnte für 16 von 77 putativen Proteinen, die unter diesen Punkt zusammengefasst wurden (Tab. 8.1), *in silico* (6.2.5) eine potentiell PPC-spezifische BTS identifiziert werden (Tab. 8.2). 12 der Kandidaten wurden über BlastP bzw. über die direkte Suche (Annotation) in der Genomdatenbank von *P. tricornutum* identifiziert, während die restlichen vier über den nicht gerichteten Ansatz (6.2.5) detektiert wurden. *In vivo* Lokalisationsstudien mit den potentiellen N-terminalen Zielsteuerungssequenzen bzw. FL-Sequenzen der Proteine fusioniert an eGFP konnten nur für vier der 16 Kandidaten die vorhergesagte PPC-

Lokalisation bestätigen (Abb. 3.17). Bei diesen handelt es sich um sPrp, sP4h, sPel und sSmc. Diese Proteine wurden hauptsächlich nach enthaltenen putativen Domänen (NCBI-CDS) benannt (Abb. 3.18).

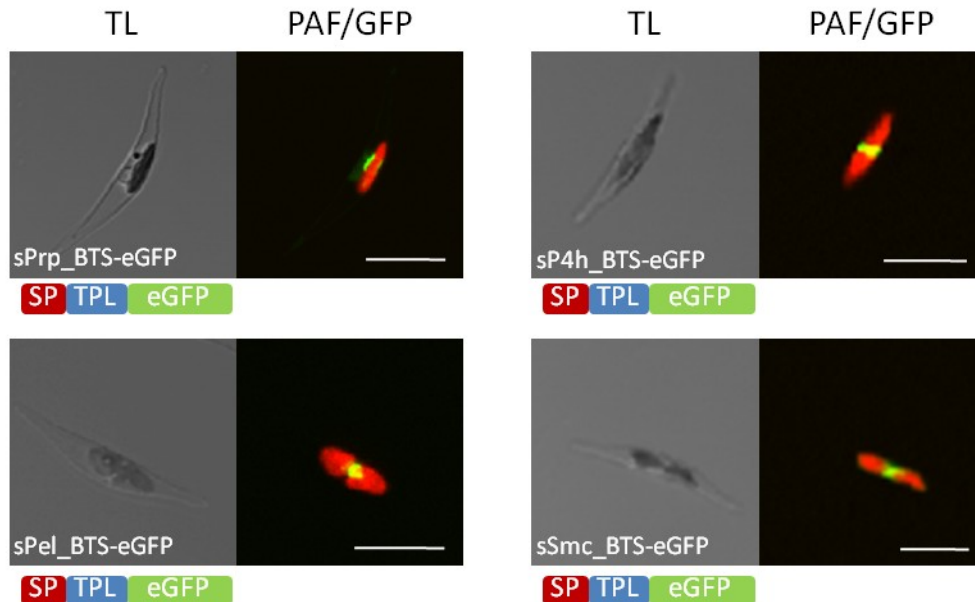


Abb. 3.17 In vivo Lokalisation von sPrp, sP4h, sPel und sSmc im PPC von *P. tricornutum*. Die Expression der BTS-Sequenzen von sPrp, sP4h, sPel und sSmc aus *P. tricornutum* fusioniert mit eGFP resultierte in einer PPC-charakteristischen Fluoreszenzstruktur der Fusionsproteine. TL, *transmitted light* (Durchlicht); PAF, Plastidenautofluoreszenz; GFP, *green fluorescent protein*; SP, Signalpeptid; TPL, *transit peptide-like sequence*. Maßstab = 10 μ m.

Für sPrp (*pentapeptide repeats containing protein*) konnten über NCBI-CDS keine konservierten Domänen identifiziert werden. Auffällig bei der Primärsequenz des Proteins (Gesamtlänge: 380 AS) war jedoch eine repetitive Abfolge der Aminosäuren T-G-P-S-L/M/A/V, welche 54 mal im Protein vorkommt, sowie ein aus 12 AS bestehendes Motiv (D-S-D-D-R-R-L-R/K-S-P-S-S), das, z.T. abgewandelt, siebenfach in der Primärsequenz vorhanden ist. Die Proteinsequenz von sPrp besteht somit außer der N-terminalen BTS, für die eine Länge von 32 AS ermittelt werden konnte (Tab. 8.2) und den letzten 5 AS am C-Terminus, ausschließlich aus repetitiven Elementen. Das beste BlastP-Ergebnis für Proteinsequenzen mit signifikanten Gemeinsamkeiten war ein als Trans-Sialidase annotiertes Protein aus *Metarhizium acridum* (Ascomycota) (*e-value*: $1e-49$). Trans-Sialidasen sind für gewöhnlich an der Biosynthese von Sialinsäuren (z.B. N-Acetylneuraminsäure) beteiligt, die eine Rolle bei der Ausbildung von Zell-Zell-Interaktionen als Oberflächendeterminanten spielen (Frasch, 2000).

Das PPC-lokalisierte sP4h (*prolyl 4-hydroxylase domain containing protein*) besitzt eine PPC-spezifische BTS mit einer abgeschätzten Länge von 135 AS, wobei das SP davon 51 AS einnimmt. Über NCBI-CDS konnte für sP4h eine konservierte Prolyl 4-Hydroxylase-Domäne, sowie zwei TPR-Motive detektiert werden (Abb. 3.18). Letztere gelten, wie bereits unter 3.1.2 erwähnt, als Domänen, die eine Rolle bei Protein-Protein-Interaktionen spielen. Prolyl 4-Hydroxylasen sind dahingegen für gewöhnlich an der posttranslationalen Modifikation von Proteinen beteiligt. Neben dem PPC-lokalisierten sP4h konnten in *P. tricornutum* zwei weitere Proteine mit Prolyl 4-Hydroxylase-Domäne *in silico* identifiziert werden, die jedoch höchstwahrscheinlich im Cytosol der Diatomee lokalisieren (Tab. 8.1).

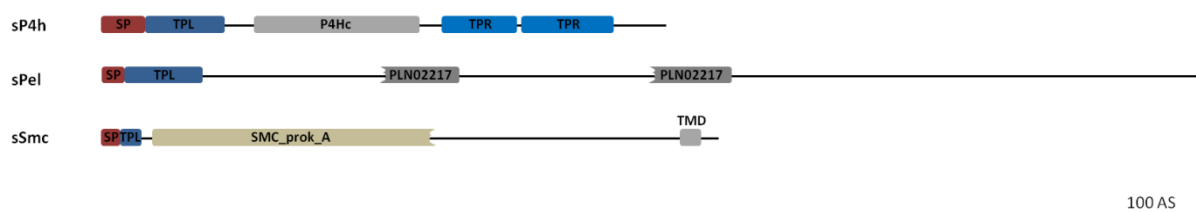


Abb. 3.18 Identifizierte Proteindomänen der PPC-lokalisierten Proteine sP4h, sPel und sSmc aus *P. tricornutum*. SP, Signalpeptid; TPL, Transitpeptid-ähnliche Sequenz; TPR, *tetratricopeptide repeats*; P4Hc, *prolyl 4-hydroxylase alpha subunit homologues*; PLN02217, *probable pectinesterase/pectinesterase inhibitor*; SMC_prok_A, *chromosome segregation protein SMC (structural maintenance of chromosomes), primarily archaeal type*; die Detektion putativer Domänen erfolgte mittels NCBI-CDS.

Bei sPel handelt es sich um ein Protein mit einer Länge von 1191 AS, wovon die N-terminale PPC-spezifische BTS ca. 109 AS einnimmt (Tab. 8.2). NCBI-CDS Analysen ergaben, dass für sPel zwei putative, jedoch gering konservierte und unvollständige Pectinesterase-/Pectinesterase Inhibitor-Domänen (PLN02217) detektiert werden konnten (Abb. 3.18). Enzyme mit solchen Domänen (z.B. Pectinmethylesterasen) sind dafür bekannt bei der Zellwandsynthese von dicotylodonen Pflanzen mitzuwirken (Micheli, 2001). BlastP-Analysen ergaben, dass ähnliche Proteine zu sPel aus anderen Organismen u.a. als PT *repeat-containing protein* (*Toxoplasma gondii*, e-value: 6e-41) oder auch Mucin (*Dictyostelium fasciculatum*, e-value: 2e-16) in der NCBI-Datenbank annotiert sind. Einer Analyse der AS-Komposition von sPel ergab, dass die Primärsequenz des Proteins zu 38,5 % aus Prolin (16,2 %) und Serin (22,3%) besteht (ProtParam). Neben sPel konnten 13 weitere Proteine mit ähnlichen Domänen in *P. tricornutum* identifiziert werden. Diese wurden jedoch basierend auf den hier angewandten Kriterien zumeist als sekretorisch/ER- oder cytosolisch-lokalisiert eingestuft (Tab. 8.1).

Das vierte und letzte PPC-lokalisierte Protein in dieser Kategorie wurde nach der Detektion putativ funktioneller Domänen als sSmc (*structural maintenance of chromosomes*) benannt. Die über NCBI-CDS identifizierte SMC-Domäne wird als unvollständig erkannt und weist eine geringe Sequenzkonservierung gegenüber SMC-Proteinen des prokaryoten Typs auf (Abb. 3.18). Gegenüber eukaryoten SMC-Proteinen, welche Funktionen bei der Chromosomen-Segregation und Zellteilung wahrnehmen, scheint mit sSmc keine Sequenzidentität vorzuherrschen. Generell finden sich in BlastP-Analysen nur sehr wenige Proteine mit homologen AS-Sequenzen in der NCBI-Datenbank (höchster e-value: 0,01). Die Länge der PPC-spezifischen BTS von sSMC beträgt ca. 40 AS. Zudem konnte über TMHMM eine putative C-terminale Transmembrandomäne (AS-Position 629-651, Gesamtlänge: 671 AS) von 22 AS vorhergesagt werden (Abb. 3.18), welche das Protein möglicherweise in der PPM oder OEM verankert.

Hinsichtlich der Lokalisation der restlichen 12 als putativ PPC-lokalisiert eingeordneten Faktoren in dieser Kategorie ergaben *in vivo* Lokalisationstudien mit eGFP-Fusionskonstrukten, dass fünf ER-, vier plastidär- und drei der Proteine mitochondriell-lokalisiert zu sein scheinen (Tab. 8.1).

3.2 Homologe zu PPC-Proteinen aus *P. tricornutum* in anderen Organismen mit komplexer Plastide rhodophytischen Ursprungs und Rotalgen

Die zunehmende Verfügbarkeit genomischer Daten von Organismen mit komplexen Plastiden, die auf ein Endosymbioseereignis mit einer Rotalge zurückgehen, und Rhodophyten ermöglicht es, bei gleichzeitiger Kenntnis über Zielsteuerungssignale und -prozesse, Kompartiment-spezifische Proteomvergleiche zwischen den Organismen anzustellen. Für die Heterokontophyte *P. tricornutum* gelang es in dieser Arbeit die Kenntnisse über das PPC-Proteom von initial 11 Proteinen (siehe 1.3) um weitere 29 auf 40 durch *in vivo* Lokalisation verifizierte Proteine zu erweitern. Über die Proteinkomposition periplastidärer Kompartimente anderer Organismen mit komplexen Plastiden rhodophytischen Ursprungs, welche von vier Membranen umgeben werden, ist jedoch vergleichsweise nur wenig bekannt.

Um zu untersuchen ob Parallelen zwischen der PPC-Proteinkomposition von *P. tricornutum* und anderen evolutionär verwandten Arten bestehen und ob sich entsprechende Homologe zu

PPC-Proteinen in Rotalgen finden, wurden *in silico* Analysen durchgeführt. Dabei wurden die Proteinsequenzen PPC-lokalisierter Proteine aus *P. tricornutum* mittels BlastP oder tBlastn gegen die Genomdatenbanken relevanter Organismen abgeglichen. Darunter *Thalassiosira pseudonana*, *Ectocarpus siliculosus* und *Aureococcus anophagefferens* als weitere Vertreter der Heterokontophyten, sowie die Haptophyte *Emiliania huxleyi*, die Cryptophyte *Guillardia theta* und die Apicomplexa *Plasmodium falciparum* und *Toxoplasma gondii*. Alle diese Organismen besitzen, wie unter 1.1.2 beschrieben, komplexe Plastiden die auf eine Rotalge zurückgehen. Vergleichend dazu wurden die vier Rotalgenspezies *Cyanidioschyzon merolae*, *Galdieria sulphuraria*, *Porphyridium cruentum* und *Calliarthron tuberculosum* untersucht, bei denen z.T. jedoch nur partielle Genomdaten verfügbar waren (*P. cruentum*, *C. tuberculosum*).

Das Screening ergab, dass insbesondere symbiontische ERAD-Faktoren (SELMA) in allen untersuchten Organismen, zusätzlich zum ERAD-System der Wirtszelle, existieren (Tab. 3.1). Vor allem zentrale Komponenten des SELMA-Systems, wie sDer1, sCdc48 und Proteine der Ubiquitinierungsmaschinerie, wurden in allen Organismen detektiert. In den analysierten Rotalgen konnten im Vergleich dazu die korrespondierenden ERAD Komponenten identifiziert werden. Faktoren die auf ein zweites (s)Cdc48, ein sPub- oder ptDUP-Homolog hindeuten, konnten in den hier untersuchten Rhodophyten nicht aufgefunden werden (Tab. 3.1). Diese scheinen in den inspizierten Apicomplexa ebenfalls zu fehlen.

In den analysierten Heterokontophyten finden sich mehr putative Homologe zu PPC-Proteinen aus *P. tricornutum* als in Haptophyten, wobei sich bei den Apicomplexa weniger als ein Viertel an putativen PPC-Homologen, hauptsächlich SELMA-Faktoren, detektieren ließen. Eine Ausnahme diesbezüglich nimmt die Cryptophyte *G. theta* ein, die ein nahezu vollständiges 26S Proteasom in ehemaligen Rotalgencytoplasma zwischen PPM und OEM besitzt (Douglas et al, 2001; Stork et al, 2012), welches z.B. bei Haptophyten und Apicomplexa auch in reduzierter Form (vgl. Heterokontophyten) nicht detektiert werden konnte (Tab. 3.1).

Tab. 3.1 Vergleich der PPC-Proteine aus *P. tricornutum* mit putativen Homologen aus anderen Organismen mit komplexer Plastide rhodophytischen Ursprungs und rezenten Rotalgen

Klasse/Funktion	Pt PPC-Protein	He				Ha	Cr	Ap		Rh				
		<i>P. tri</i>	<i>T. pse</i>	<i>E. sil</i>	<i>A. ano</i>	<i>E. hux</i>	<i>G. the</i>	<i>P. fal</i>	<i>T. gon</i>	<i>C. mer</i>	<i>G. sul</i>	<i>P. cru</i>	<i>C. tub</i>	
SELMA	sDer1-1													
	sDer1-2													
	ptE3P													
	sUba1													
	sUbc4						4							
	sUbc6													
	sUb													
	sCdc48-1													
	sCdc48-2													
	sUfd1													
	sPub													
	ptDUP													
Chaperone	sHsp70													
	sDpc													
	sDtc													
	sDnaK													
Proteasomale 20S-Komponenten	sAlpha1													
	sAlpha2													
	sAlpha3	2	2											
	sAlpha4													
	sAlpha5													
	sAlpha6													
	sAlpha7	2												
	sAlpha													
	sBeta1													
	sBeta2													
	sBeta3													
	sBeta4													
	sBeta5													
	sBeta6													
sBeta7														
Proteasen	sTlp1													
	sTlp2													
	sTlp3													
	sTlp4													
	sTlp5		6			4						2	3	
	sTlp6													
Lipidtransfer	sSec14													
Plastidenteilung	sDrp													
	sPdr1				2									
<i>G. theta</i> Nucleomorph-Homologe	sORF532a													
	sORF534													
	sORF261												2	
	sORF139													
Unbekannt	sSmc													
	sPel													
	sP4h													
	sPrp													
Metabolismus	sαCa1						2							
	sαCa2													
	s6PGDH													
Regulation	sTrxH					2	2							
	sNTRC													

He, Heterokontophyta; Ha, Haptophyta, Cr, Cryptophyta; Ap, Apicomplexa; Rh, Rhodophyta; *P. tri*, *Phaeodactylum tricornutum*; *T. pse*, *Thalassiosira pseudonana*; *E. sil*, *Ectocarpus siliculosus*; *A. ano*, *Aureococcus anophagefferens*; *E. hux*, *Emiliana huxleyi*; *G. the*, *Guillardia theta*; *P. fal*, *Plasmodium falciparum*; *T. gon*, *Toxoplasma gondii*; *C. mer*, *Cyanidioschyzon merolae*; *P. cru*, *Porphyridium cruentum*; *C. tub*, *Calliarthron tuberculosum*. Grün ausgefüllte Felder deuten an, dass putative Homologe bzw. Orthologe

über BlastP/tBlastn in der Genomdatenbank der jeweiligen Spezies (bzw. der NCBI-Datenbank) mit potentiell PPC-spezifischer BTS (nicht bei Rhodophyten) identifiziert werden konnten. Zahlen verweisen auf die Anzahl detektierter Homologe. Im Fall von leeren Feldern wurde kein potentielles Homolog detektiert. Zur näheren Erläuterung, Angabe der Protein-IDs und Quellen siehe 8.5.

Neben dem SELMA-System konnten, mit Ausnahme von sORF139, Homologe zu den symbiontischen ORFs aus *P. tricornutum* (ursprünglich als Homologe zu Nucleomorph-kodierten ORFs aus *G. theta* identifiziert) weitestgehend in den hier analysierten sekundär evolvierten Spezies und Rotalgen identifiziert werden (Tab. 3.1). Gleiches traf mit Ausnahmen auch für das PPC-ansässige Thioredoxin-System und den Plastidenteilungsapparat (cytosolisch in Rhodophyten) zu. Die *in silico* Analysen ergaben im Weiteren, dass sich potentiell PPC-lokalisierte Homologe zu Komponenten wie sTlp1-6, sSmc, sPel, sP4h, sPrp, sowie s6PGDH nur partiell in den untersuchten Organismen mit komplexen Plastiden rhodophytischen Ursprungs und analysierten Rotalgenspezies auffinden ließen (Tab. 3.1).

3.3 Putative Triosephosphat/Phosphat-Translokatoren in *P. tricornutum*

Phototrophe Eukaryoten profitieren metabolisch enorm von ihren akquirierten Endosymbionten (primäre oder komplexe Plastiden). Letztere versorgen die Wirtszelle mit energiereichen Metaboliten, welche innerhalb der Photosynthese durch Fixierung von Kohlenstoff synthetisiert werden (Calvin-Zyklus) (Flügge, 1999). Der Austausch von Metaboliten zwischen Plastide und Wirtscytosol erfolgt für gewöhnlich über spezifische Translokatoren (TPTs, PPTs, GPTs, XPTs), die sich bei primären Plastiden in der inneren Plastidenmembran (IEM) befinden und Substrate über einen Antiportprozess über die Membran transportieren (Joyard et al, 1998). Komplexe Plastiden besitzen jedoch eine oder zwei zusätzliche Hüllmembranen, was die Präsenz weiterer Transportkomponenten verlangt.

Gene kodierend für putative TPTs konnten in *P. tricornutum* zwar bereits identifiziert werden, über die Lokalisation der Genprodukte ist jedoch nur sehr wenig bekannt. Um die Präsenz putativer TPTs in den Membranen der komplexen Plastide von *P. tricornutum* zu untersuchen wurden bioinformatische Analysen, kombiniert mit *in vivo* Lokalisationsstudien durchgeführt. Der Ergebnisse dieser Studien sind in den folgenden Abschnitten beschrieben.

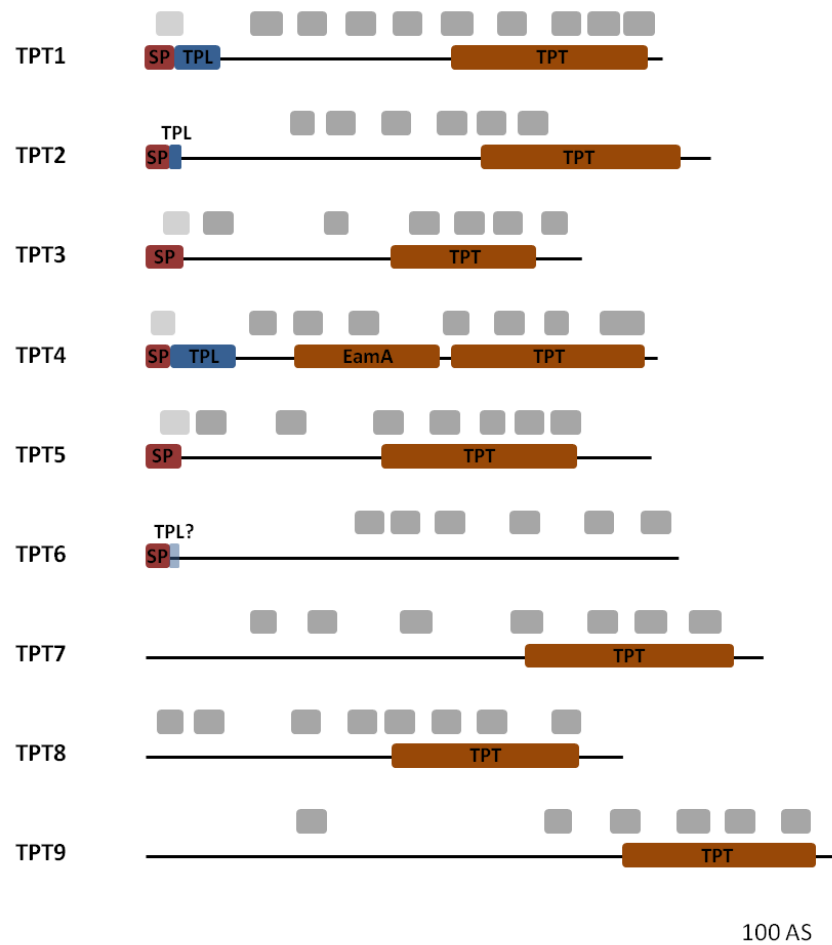


Abb. 3.19 Domänen-Alignment putativer Triosephosphat/Phosphat-Translokatoren aus *P. tricornutum*. Mit Ausnahme von TPT6 wurde für alle Proteine durch NCBI-CDS eine konservierte TPT-Domäne (*triosephosphate transporter family*) erkannt. Für TPT4 wurde zusätzlich eine EamA-Domäne (*hypothetical membrane protein of unknown function*) detektiert. Durch Analyse der Primärsequenzen der Proteine mit SignalP konnte für TPT1 bis TPT6 ein SP vorhergesagt werden. Zusätzliche putative TPL-Sequenzen wurden für TPT1 (plastidär), TPT2 (PPC bzw. PPM), TPT4 (plastidär) und TPT6 (plastidär) ermittelt (siehe Tab. 3.2). Für die putativen TPTs wurden durch TMHMM jeweils 6 bis 10 Transmembrandomänen (TMDs) vorhergesagt (in grau). TMDs in hellgrau deuten darauf hin, dass das potentielle TMD-Segment innerhalb der vorhergesagten N-terminalen Zielsteuerungssequenz lokalisiert ist und wahrscheinlich den hydrophoben Bereich des SP wiedergibt (TMD unwahrscheinlich), zumal für keines der Proteine eine Signalankersequenz vorhergesagt wurde.

Eine Präsequenzanalyse mit SignalP ergab, dass für TPT1 bis TPT6 potentielle Signalpeptide ermittelt werden konnten (Tab. 3.2). Für die restlichen TPTs (TPT7-9) konnte dahingegen keine N-terminale Zielsteuerungssequenz identifiziert werden. Während die Sequenzen von TPT3 und TPT5 direkt hinter dem vorhergesagten SP konserviert sind und somit eine weitere TPL-Sequenz unwahrscheinlich ist, besitzen TPT1, TPT2, TPT4 und TPT6 neben dem SP erweiterte unkonservierte N-terminale Bereiche von 35 (TPT1), 9 (TPT2), 49 (TPT4) und 5 (TPT6) Aminosäuren Länge, welche jeweils für eine TPL-Sequenz in Frage kommen. Die +1-Position nach der potentiellen SP-Schnittstelle wird bei TPT1, TPT4 und TPT6, zumeist im Kontext eines „ASA-FAP-Motivs“ (plastidäres *targeting*), von einer aromatischen AS oder Leucin eingenommen. Dahingegen findet sich bei TPT2 dort die AS Asparaginsäure (Tab.3.2).

Tab. 3.2 Detektierte Präsequenzen putativer TPTs aus *P. tricornutum*

Protein	Putative Präsequenz
TPT1_BTS	<u>MMMKRALVVLTL</u> <u>SVGVSAR</u> ASA FAPGAAVKNHAGATQSAIHKQTPFPPTTELEKLRPQTSAL...
TPT2_BTS	<u>MRFAAWLVILTGT</u> FVEADRVQSTQPQSASN...
TPT3_SP	<u>MSSSTAKDSGAASPLKLFVLVVCWYAGNT</u> FYNIY...
TPT4_BTS	<u>MKVATTLTLAFIC</u> ASA FGLNGQTTSVMKKVGF DAGSKPMVQAIDVQGNRLGSNMQPLKSAVA NEDAPRGGAT...
TPT5_SP	<u>MAQTKASQSSTLWLLVWMVNNIGVTLL</u> NKAAF...
TPT6_BTS?	<u>MLVIAILLTFV</u> VH STA FPPTVSFGSS...

Fett, SP nach HMM (SignalP 3.0); unterstrichen, SP nach NN (SignalP 3.0); normal, putative TPL-Sequenz (N-terminale Extension nach BlastP); kursiv, konserviertes matures Protein; in Rot, potentielles „ASA-FAP Motiv“ (plastidäres *targeting*). Für *in vivo* Lokalisationsstudien mit Präsequenz-eGFP Fusionskonstrukten wurden z.T. längere Sequenzbereiche exprimiert um die Integrität der N-terminalen Zielsteuerungssequenzen zu gewährleisten.

Blast-Analysen ergaben, dass sich für TPT1-4 vor allem Homologe in Organismen mit komplexen Plastiden rhodophytischen Ursprungs und Rotalgen finden und diese zumeist als Triosephosphat/Phosphat-Translokatoren annotiert sind. Für TPT5 ließen sich, mit Ausnahme homologer Proteine in Heterokontophyten (Oomyceten), hauptsächlich Homologe aus Pilzen in der NCBI-Datenbank detektieren. TPT6, TPT7 und TPT9 besitzen nach BlastP insbesondere Homologe in Pflanzen und Grünalgen, die u.a. als Triosephosphat oder Phosphoenolpyruvat/Phosphat-Translokatoren (TPT6 und TPT7) annotiert sind. Homologe zu

TPT8 sind dahingegen vorwiegend als GDP-Fucose Transporter (Golgi Apparat) annotiert und finden sich hauptsächlich in den Metazoa.

Die Genmodelle von TPT1(54017), TPT2 (20826), TPT3 (55414), TPT4 (8738) und TPT9 (50485) sind vollständig mit EST-Daten abgedeckt, während für TPT5 (40415), TPT6 (32506), TPT7 (34918) und TPT8 (43174) nur partielle EST-Sequenzen aufzufinden waren. Für *in vivo* Lokalisationsstudien wurden alle TPT-kodierenden Sequenzen aus cDNA amplifiziert, wobei das Genmodell von TPT8 am 3'-Ende im Bereich eines vorhergesagten Introns korrigiert wurde. Um die Lokalisation der neun putativen TPTs *in vivo* zu bestimmen wurden GFP-Lokalisationsstudien sowohl mit den detektierten N-terminalen Zielsteuerungssequenzen als auch mit FL-Proteinen durchgeführt. Während für die Präsequenzen von TPT3 und TPT6 fusioniert mit eGFP eine Lokalisation im (c)ER beobachtet werden konnte, resultierte die Expression der Präsequenz von TPT2 zusammen mit eGFP in einer PPC-typischen „Blob-ähnlichen“ Fluoreszenzstruktur (BLS). Die BTS-Sequenzen von TPT4 und TPT1 dirigierten dahingegen eGFP in das Stroma der Plastide (Abb. 3.20A). Im Fall von TPT4 konnte dies bereits in vorangehenden Studien beobachtet werden (Kilian & Kroth, 2004; Kilian & Kroth, 2005).

Die Resultate der Expression von FL-eGFP Fusionskonstrukten der TPTs waren mit denen der Präsequenzen weitestgehend kongruent. So zeigte sich für TPT3_FL-eGFP eine Lokalisation im (c)ER bzw. der (c)ER-Membran, wobei die FL-Sequenz von TPT6 fusioniert mit eGFP nicht wie bei TPT3 im (c)ER oder der (c)ER-Membran sondern eher dem PPC bzw. der PPM zu lokalisieren schien (BLS; Abb. 3.20B). Letztere Lokalisation konnte auch für TPT2_FL-eGFP beobachtet werden und ist im Einklang mit der Lokalisation des Präsequenz-eGFP Fusionsproteins im PPC.

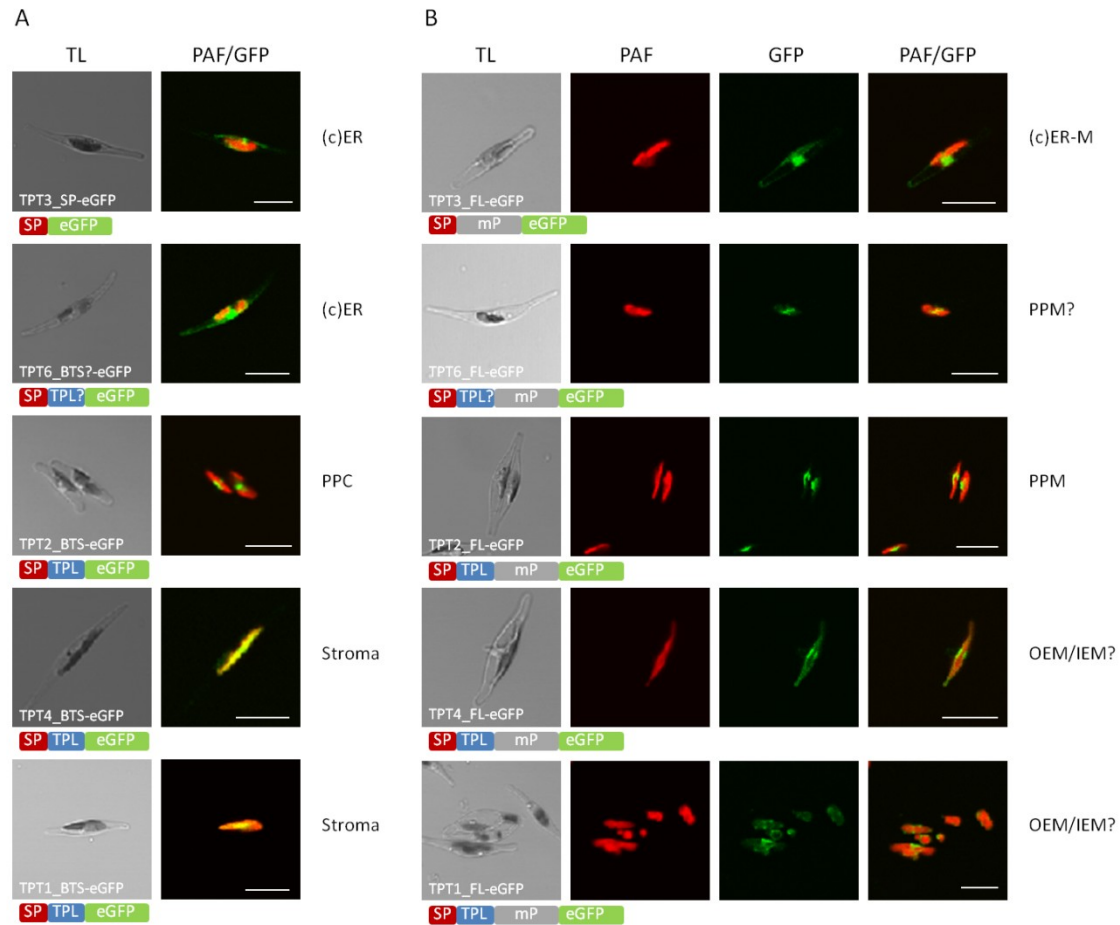


Abb. 3.20 In vivo Lokalisation putativer TPTs in der komplexen Plastide von *P. tricornutum*. (A) Expression der TPT-Präsequenzen als Fusion mit eGFP. (B) Lokalisation der FL-TPT-Sequenzen fusioniert mit eGFP. Die Expression der Fusionskonstrukte erfolgte über den lichtinduzierbaren pPha-T1 Vektor. ER, endoplasmatisches Retikulum; (c)ER, (*chloroplast*) ER; cER-M, cER-Membran; PPC, periplastidäres Kompartiment; PPM, periplastidäre Membran; OEM/IEM, *outer/inner chloroplast envelope membrane*; TL, *transmitted light* (Durchlicht); PAF, Plastidenautofluoreszenz; GFP, *green fluorescent protein*; SP, Signalpeptid; TPL, *transit peptide-like sequence*; mP, matures Protein. Maßstab = 10 μ m.

Die Expression von TPT4 und TPT1 als FL-eGFP Fusionskonstrukte führte, wie Abbildung 3.20B verdeutlicht, zu einer GFP-Fluoreszenzstruktur die sich teilweise punktförmig in den Randbereichen der Plastidenautofluoreszenz konzentriert bzw. diese umrandet. Basierend auf diesen Beobachtungen konnte keine Differenzierung getroffen werden, ob es sich bei der Lokalisation von TPT4 bzw. TPT1 um eine Lokalisierung in der OEM oder IEM der komplexen Plastide handelt.

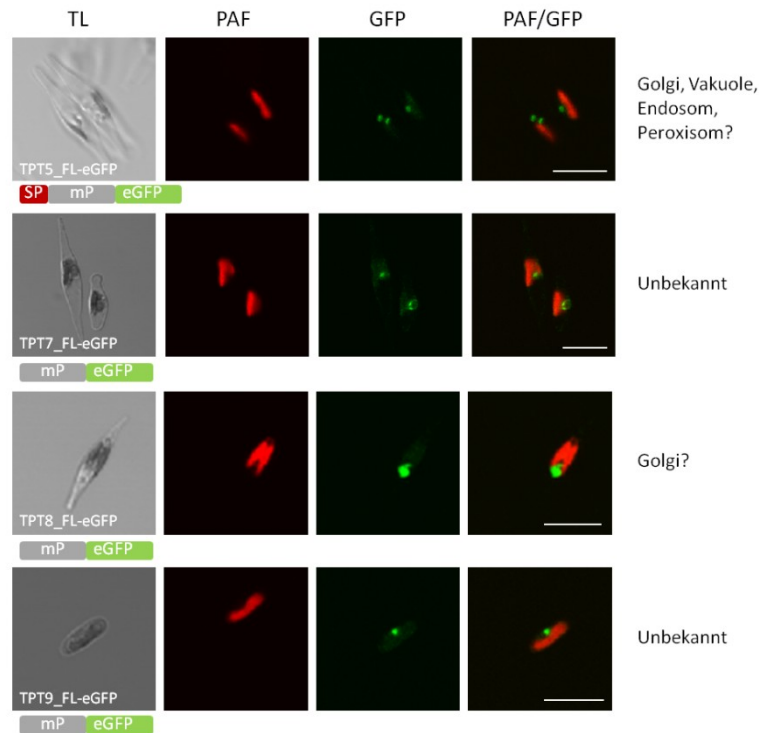


Abb. 3.21 *In vivo* Lokalisation weiterer TPTs außerhalb der komplexen Plastide von *P. tricornutum*. Die Expression von FL-eGFP Fusionskonstrukten mit TPT5, TPT7, TPT8 und TPT9 führten jeweils zu einer Lokalisation außerhalb der komplexen Plastide. TL, *transmitted light* (Durchlicht); PAF, Plastidenautofluoreszenz; GFP, *green fluorescent protein*; SP, Signalpeptid; mP, matures Protein. Maßstab = 10 μ m.

In vivo Lokalisationsstudien mit TPT5, TPT7, TPT8 und TPT9 ergaben, dass diese vier Proteine offensichtlich nicht in Membranen der komplexen Plastide von *P. tricornutum* lokalisieren (Abb. 3.21). Alle vier Fusionsproteine lokalisierten in konzentrierten punktartigen Strukturen, die sich klar von der Autofluoreszenz der Plastide abgrenzen.

3.3.2 Untersuchung der Zielmembranen und Orientierungen putativer TPTs in der komplexen Plastide von *P. tricornutum* über das Self-Assembling-GFP System

Die unter 3.3.1 gemachten Beobachtungen deuten auf die Präsenz von mindestens 4 putativen TPTs (TPT1-4, TPT6) in der komplexen Plastide von *P. tricornutum* hin. Allerdings konnte anhand der *in vivo* Lokalisationsstudien die genaue Lokalisation von TPT6 (PPM?), sowie TPT1 und TPT4 (plastidäre OEM/IEM?) nicht geklärt werden. Außerdem stellt sich die Frage nach der Orientierung der putativen TPTs in den jeweiligen Zielmembranen innerhalb der komplexen Plastide. Diesen Fragen wurde daher im Weiteren mittels Anwendung des Self-Assembling-GFP Systems (SA-GFP) nachgegangen (6.2.3.10.2). Die spezifischen

Markerproteine bzw. Zielsteuerungssequenzen vor GFP_S1-10 für die Subkompartimente der komplexen Plastide von *P. tricornutum*, welche in dieser Studie zum Einsatz kamen waren: GFP_S1-10 (Cytosol); PDI_FL (ER); sHsp70_BTS (PPC); MGD1_FL (Intermembranraum) und AtpC_BTS (Plastidenstroma) (siehe 6.2.3.10.2).

Die putativen TPTs der komplexen Plastide von *P. tricornutum* wurden dabei als Fusionsproteine mit GFP_S11 (C-terminal) jeweils zusammen mit einem spezifischen Markerprotein fusioniert mit GFP_S1-10 (C-terminal) aus einem angrenzenden Kompartiment der potentiellen Zielmembran exprimiert. Die duale Expression erfolgte über den Vektor pPha_Dual_2xNR für sechs Stunden nach Induktion mit Nitrat (6.2.3.10.2), vor Betrachtung der Klone unter dem KLSM (6.2.4). Wie Abbildung 3.22 verdeutlicht, konnte über SA-GFP für TPT3_FL_S11 eine Lokalisation in der (c)ER-Membran bestätigt werden (siehe 3.3.1). Nach Expression des TPT3-Fusionsproteins mit cytosolischem GFP_S1-10 konnte eine ER-typische GFP-Fluoreszenz beobachtet werden, wobei die Expression mit dem ER-Marker fusioniert mit GFP_S1-10 zu keiner Fluoreszenz führte. Der C-Terminus von TPT3 ist dem Cytosol zugewandt. Im Fall von TPT6, dessen putative Präsequenz eGFP in das ER dirigierte und bei dem die Expression der FL-Sequenz in einer möglichen PPM-Lokalisation resultierte (Abb. 3.20), ergab sowohl die gleichzeitige Expression von TPT6_FL_S11 mit dem ER-Marker als auch mit dem PPC-Marker fusioniert mit GFP_S1-10 GFP-Fluoreszenz. Zielmembran ((c)ER-Membran oder PPM) und Orientierung des C-Terminus von TPT6 konnten durch diesen Ansatz nicht eindeutig bestimmt werden.

Bezüglich der Lokalisation von TPT2 konnte für die Präsequenz, sowie das FL-Protein fusioniert mit eGFP bereits unter 3.3.1 eine relativ eindeutige PPC- bzw. PPM-Lokalisation beobachtet werden. Allerdings konnte die Orientierung des Proteins in der PPM über SA-GFP nicht definitiv geklärt werden. Eine duale Expression von TPT2_FL_S11 sowohl mit dem ER-spezifischen als auch dem PPC-spezifischen Marker fusioniert mit GFP_S1-10 resultierte jeweils in GFP-Fluoreszenz (Abb. 3.23). Dabei führte die Expression mit dem ER-Marker zu einer typischen BLS, wobei sich die Fluoreszenzstruktur bei Expression mit dem PPC-Marker nicht nur auf den ansonsten typischen Bereich zwischen den beiden Plastidenhälften beschränkte, sondern punktiert erweitert war.


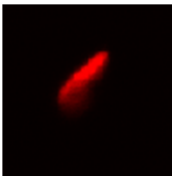
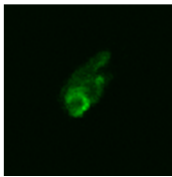
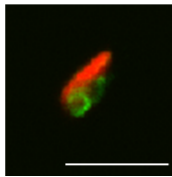
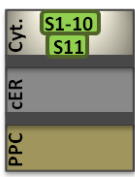
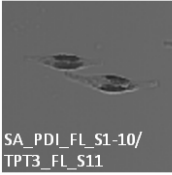
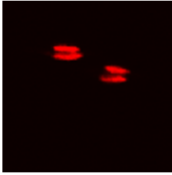

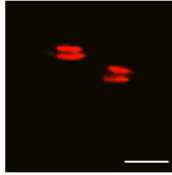
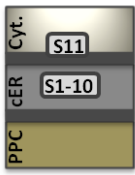

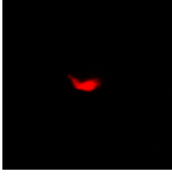
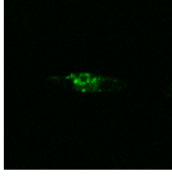
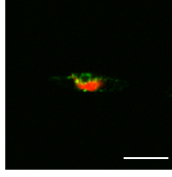



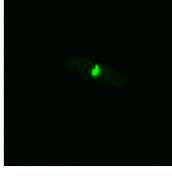
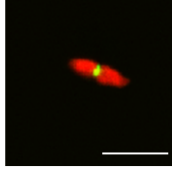


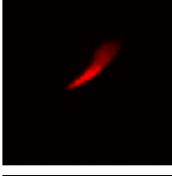
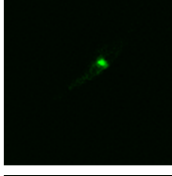
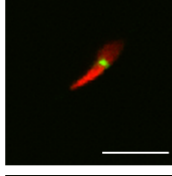


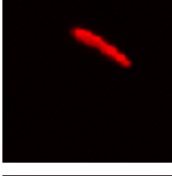
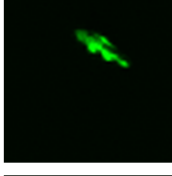
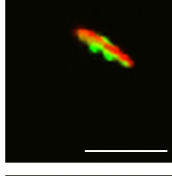
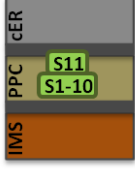

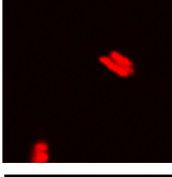
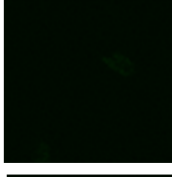
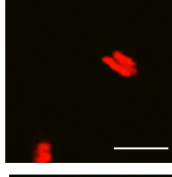
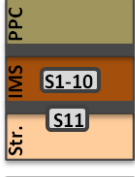



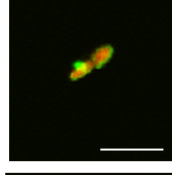
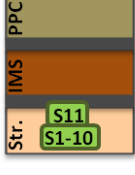

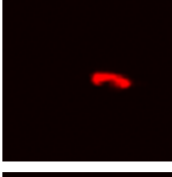

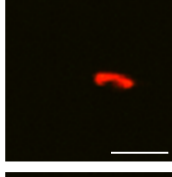
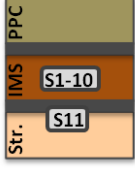

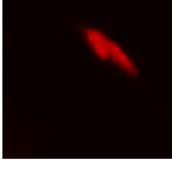
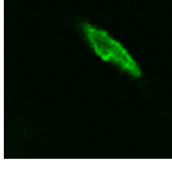
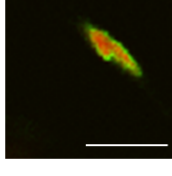
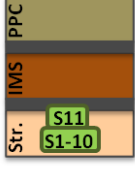
		TL	PAF	GFP	PAF/GFP		
TPT3	Cytosol Marker	 SA_S1-10/ TPT3_FL_S11					(c)ER-M, C-Term. Cytosol
	ER Marker	 SA_PDI_FL_S1-10/ TPT3_FL_S11					
TPT6	ER Marker	 SA_PDI_FL_S1-10/ TPT6_FL_S11					(c)ER-M/PPM?, C-Term.?
	PPC Marker	 SA_sHsp70_BTS_S1-10/ TPT6_FL_S11					
TPT2	ER Marker	 SA_PDI_FL_S1-10/ TPT2_FL_S11					PPM, C-Term. PPC?
	PPC Marker	 SA_sHsp70_BTS_S1-10/ TPT2_FL_S11					
TPT4	IMS Marker	 SA_MGD1_FL_S1-10/ TPT4_FL_S11					IEM, C-Term. Stroma
	Stroma Marker	 SA_AtpC_BTS_S1-10/ TPT4_FL_S11					
TPT1	IMS Marker	 SA_MGD1_FL_S1-10/ TPT1_FL_S11					IEM, C-Term. Stroma
	Stroma Marker	 SA_AtpC_BTS_S1-10/ TPT1_FL_S11					

Abb. 3.22 Analyse der Lokalisation und Membranorientierung putativer TPTs in der komplexen Plastide von *P. tricornutum* mit dem Self-Assembling-GFP System. Die FL-Sequenzen von TPT1, TPT2, TPT3, TPT4 und TPT6 wurden jeweils als Fusionsproteine mit C-terminalem GFP_S11 zusammen mit einem spezifischen Markerprotein eines potentiell an die Zielmembran angrenzenden Kompartiments fusioniert an GFP_S1-10 exprimiert. Cyt., Cytosol; C-Term., C-Terminus; ER, endoplasmatisches Retikulum; (c)ER-M, (*chloroplast*) ER-Membran; PPC, periplastidäres Kompartiment; PPM, periplastidäre Membran; IMS, Intermembranraum; OEM/IEM, *outer/inner chloroplast envelope membrane*; TL, *transmitted light* (Durchlicht); PAF, Plastidenautofluoreszenz; GFP, *green fluorescent protein*. Maßstab = 10 µm.

Abschließend wurden die beiden plastidären TPTs, TPT1 und TPT4, mittels SA-GFP untersucht. Die Analysen ergaben, dass die jeweils gleichzeitige Expression beider Proteine (Fusion mit GFP_S11) mit dem ins Plastidenstroma dirigierten AtpC_BTS-GFP_S1-10 Fusionsprotein in einer die Plastidenautofluoreszenz umrandenden GFP-Fluoreszenz resultierte (Abb. 3.22) wohingegen es bei Expression mit dem IMS-Marker zu keiner Fluoreszenzausbildung kam.

3.3.3 Analyse der Membranintegration putativer TPTs in den Hüllmembranen der komplexen Plastide von *P. tricornutum*

Nachdem für fünf putative TPTs (TPT1-4, TPT6) eine Lokalisation in der komplexen Plastide von *P. tricornutum* mit großer Wahrscheinlichkeit bestätigt werden konnte (3.4.1/2), sollte abschließend überprüft werden ob diese auch tatsächlich in die Membranen des Endosymbionten inseriert werden. Die Analysen wurden hier auf jene TPTs mit putativ rhodophytischem Ursprung, d.h. für die Homologe in Rotalgen und verwandten sekundär evolvierten Organismen detektiert werden konnten, beschränkt (TPT1-4). Die Überprüfung einer potentiellen Membraninserierung von TPT1-4 erfolgte über die Methode der Carbonatextraktion (6.2.3.7). Dazu wurde für jeden der vier TPTs zunächst ein Expressionstest durchgeführt, bei dem jeweils vier unterschiedliche TPT_FL-eGFP exprimierende Klone hinsichtlich ihres Expressionslevels untersucht wurden (Abb. 8.5). Diejenigen Klone, bei denen die höchste Expression an Fusionsprotein beobachtet werden konnte (Abb. 8.5), wurden im Anschluss zur Carbonatextraktion eingesetzt.

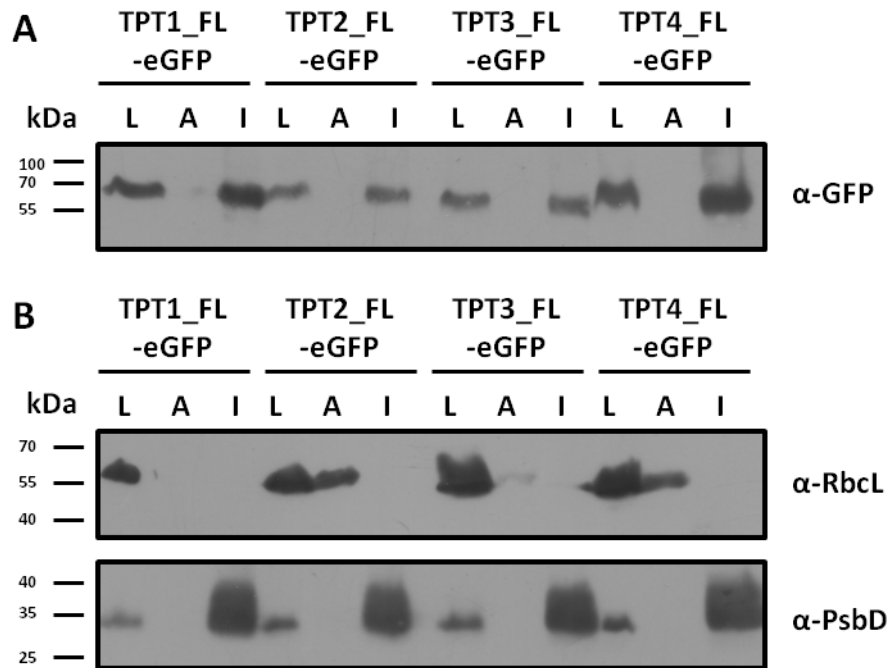


Abb. 3.23 Western Blot nach Carbonatextraktion mit den TPTs der komplexen Plastide von *P. tricornutum*. Von allen Fraktionen (L, A, I) wurden gleiche Verhältnisse aufgetragen (15 μ l). (A) Die Detektion von TPT1-4_FL-eGFP Fusionsproteinen erfolgte mittels GFP-Antikörper (1:3000). (B) Als Kontrollen wurden Antikörper gegen die große Untereinheit von RuBisCO RbcL (löslich) und eine Untereinheit des Photosystem II PsbD (integral) eingesetzt 6.1.6). L, lösliche Proteinfraction; A, assoziierte Proteinfraction; I, integrale Proteinfraction; kDa, Kilodalton. Berechnete molekulare Massen (PeptideMass): TPT1_FL-eGFP: 68,5 kDa, TPT2_FL-eGFP: 73,8 kDa, TPT3_FL-eGFP: 62,5 kDa, TPT4_FL-eGFP: 68,8 kDa.

Wie Abbildung 3.23 verdeutlicht, konnten die TPT_FL-eGFP Fusionsproteine (TPT1-4) hauptsächlich in der integralen Proteinfraction detektiert werden, wobei jedoch auch ein Anteil in der löslichen Proteinfraction vorhanden war. Verglichen mit der Kontrolle für die integrale Proteinfraction, PsbD (Photosystem II, Thylakoidmembran), zeigte sich dort ebenfalls, dass ein geringer Anteil an Protein zusätzlich zur integralen in der löslichen Proteinfraction detektiert werden konnte. Bei der Kontrolle für lösliche Proteine, der großen Untereinheit von RuBisCO RbcL, erschienen vor allem bei Klonen von TPT2_FL-eGFP und TPT4_FL-eGFP zusätzlich zu starken Signalen in der löslichen auch Signale in der assoziierten Proteinfraction.

4. Diskussion

4.1 Neue Proteine und Funktionen, sowie abwesende Prozesse im PPC von *P. tricornutum*

Diatomeen sind ökologisch höchst relevante Organismen und besitzen komplexe Plastiden mit vier Hüllmembranen, welche auf ein Endosymbioseereignis mit einer Rotalge zurückgehen (Armbrust, 2009; Lang et al, 1998). Der rhodophytische Endosymbiont wurde im Laufe der Evolution stark reduziert, wobei essentielle Gene in den Nucleus der Wirtszelle transferiert wurden und der ursprüngliche Nucleus der Rotalge, sowie sonstige Organellen (außer der primären Plastide) im ehemaligen Cytoplasma der internalisierten Rotalge (periplastidäres Kompartiment, PPC) schließlich verloren gingen (Keeling, 2010). Trotz enormer Minimierung erhalten Diatomeen das PPC als Subkompartiment ihrer komplexen Plastide, wobei über die Funktion dieses Kompartiments bisher nur wenig bekannt ist. Dort ansässige Proteine sind in Diatomeen im Nucleus der Wirtszelle kodiert und müssen durch Translokation über die beiden äußeren Membranen der komplexen Plastide (cER-Membran und PPM) ins PPC importiert werden. In der Diatomee *P. tricornutum* belaufen sich die Kenntnisse über das PPC bislang auf die Präsenz einer ERAD-ähnlichen Proteintranslokationsmaschinerie über die PPM (SELMA), sowie zwei Redox-regulatorische Proteine, ein Chaperon und ein Enzym des oxidativen Pentosephosphatwegs (Gould et al, 2006b; Gruber et al, 2009; Hempel et al, 2010; Sommer et al, 2007; Weber et al, 2009).

Um weitere Erkenntnisse über den Erhalt bzw. die Funktion des PPC zu erlangen, wurde die Proteinkomposition dieses minimierten eukaryoten Cytoplasmas im Modellorganismus *P. tricornutum* untersucht. Dabei gelang es im Rahmen dieser Arbeit über einen kombinierten *in silico/in vivo* Ansatz 29 neue PPC-Proteine in *P. tricornutum* zu identifizieren. Darunter u.a. weitere potentielle SELMA-Faktoren, proteasomale 20S Komponenten und Proteasen, Chaperone, Plastidenteilungsfaktoren, metabolisch relevante Proteine und mögliche strukturelle Komponenten (3.1). In den folgenden Abschnitten werden die putativen Funktionen, die sich aufgrund der Präsenz dieser Proteine im PPC ergeben, sowie die Konsequenz scheinbar abwesender Prozesse und Faktoren im PPC der Diatomee diskutiert.

4.1.1 SELMA und faltungsrelevante Faktoren

Cdc48-Proteine sind für ihre Beteiligung an zahlreichen diversen zellulären Vorgängen wie z.B. homotypischer Membranfusion und Zellzyklus bekannt (Latterich et al, 1995; Meyer et al, 2012). Vor allem aber spielen sie eine zentrale Rolle bei der Ubiquitin-abhängigen Dislokation und Degradation fehlerhaft gefalteter Proteine im Rahmen der ER-assoziierten Degradation (Wolf & Stolz, 2011). Beim Transport von nucleuskodierten Proteinen über die zweite Membran (PPM) der komplexen Plastide von *P. tricornutum* im Rahmen von SELMA, wird vermutet, dass sCdc48, zusammen mit dem putativen Cofaktor sUfd1, den energieabhängigen Import ubiquitinierten Substrate ins PPC vollzieht (Sommer et al, 2007). Da im Rahmen dieser Arbeit ein zweites periplastidäres sCdc48 (sCdc48-2) in *P. tricornutum* identifiziert werden konnte (3.1.1), stellt sich die Frage nach der Relevanz von zwei sCdc48-Proteinen im PPC der Diatomee.

Es ist bekannt, dass sich Cdc48-Proteine für gewöhnlich in homohexameren Strukturen organisieren (Dreveny et al, 2004), Interaktionsstudien mit dem Split-GFP-System (Kerppola, 2006) ergaben jedoch, dass sCdc48-1 mit sCdc48-2 im PPC interagiert (3.1.1.1), während eine homotypische Interaktion von sCdc48-1 nicht nachgewiesen werden konnte (Franziska Hempel, persönliche Kommunikation). Diesen Ergebnissen zu Folge scheint zwischen sCdc48-1 und sCdc48-2 eine Heterooligomerisierung stattzufinden. Es ist jedoch zudem möglich, dass es sich bei der beobachteten sCdc48-1/2-Interaktion um eine transiente Form der Proteininteraktion handelt, welche ebenfalls potentiell über das Split-GFP-System detektiert werden kann (Kerppola, 2006). Eine Homooligomerisierung von sCdc48-1 oder sCdc48-2 im PPC kann nicht ausgeschlossen werden, da der Nachweis stark vom experimentellen Aufbau des Split-GFP-Systems abhängig ist. Mögliche Störfaktoren dabei sind die Positionierung der fusionierten N- bzw. C-GFP Sequenzen an den Interaktionspartnern (N- oder C-terminal), sowie die Faltung des gesamten Fusionsproteins, welche einen entscheidenden Einfluss auf das Ergebnis des angewandten Systems haben können (räumliche Nähe von N- und C-GFP). In solchen Fällen könnten zusätzliche Interaktionsstudien über *Pull-Down-Assays* dienlich sein um potentielle Proteininteraktionen zu untersuchen.

Die Interaktion der beiden symbiontischen Cdc48-Proteine im PPC von *P. tricornutum* lässt vermuten, dass beide Proteine in Prozesse des Proteinimports über die PPM im Rahmen von SELMA involviert sind (Abb. 4.2). Es ist jedoch ebenfalls möglich, dass die Interaktion nur

transient im Zuge des Importprozesses über die PPM, bei dem sCdc48 als zentrale, energieliefernde Komponente gilt (Sommer et al, 2007), stattfindet und nur eines der Proteine tatsächlich am Importprozess mitwirkt, wobei das andere möglicherweise eine abweichende Funktion ausübt (siehe 4.1.2). Die genetische Manipulation der Diatomee über *knockdown*-Experimente eines der beiden Proteine könnte diese Frage vermutlich klären (De Riso et al, 2009). Ein zweites sCdc48 konnte, mit Ausnahme der Apicomplexa, auch in allen anderen untersuchten Organismen mit komplexen Plastiden rhodophytischen Ursprungs identifiziert werden und scheint somit eine konservierte Rolle im PPC zu spielen (siehe 3.2).

Mit sPub konnte ein weiteres PPC-lokalisiertes Protein in *P. tricornutum* nachgewiesen werden (3.1.1), für das eine Beteiligung an SELMA bzw. dem Importprozess von Präproteinen ins PPC in Frage kommt (Abb. 4.2). Neben einer Thioredoxin-Domäne, welche u.a. bei Redox-regulatorischen Proteinen vorkommt und z.B. zur Ausbildung von Disulfidbrücken an Zielproteinen dient (Holmgren, 1985), besitzt sPub eine PUB-Domäne (3.1.1). Letztere findet sich zumeist bei Proteinen mit enger Assoziation an das Ubiquitin-/Proteasom-System (Madsen et al, 2009; Suzuki et al, 2001). Zu diesen gehören u.a. PNGase-UBX-, oder UBA-Domänen beinhaltende Proteine (Allen et al, 2006), welche mit dem Prozess der ER-assoziierten Degradation (ERAD) in Verbindung gebracht werden können. PUB-Domänen gelten darüber hinaus als Cdc48-Interaktionsmodule, wobei diese vor allem mit dem C-Terminus von Cdc48-Proteinen, der als Andockstelle für substratprozessierende Faktoren gilt, zu interagieren scheinen (Allen et al, 2006; Madsen et al, 2009; Zhao et al, 2007). In diesem Kontext konnte in einer Studie mit der PUB-Domäne der PNGase (Deglycosylierung von ERAD-Substraten) aus *Mus musculus* gezeigt werden, dass die Interaktion dieser Domäne mit Cdc48 über ein Interaktionsmotiv in den letzten 10 Aminosäuren von Cdc48 stattfindet (Zhao et al, 2007). Dabei scheinen die konservierten Aminosäuren DDDLYS von Cdc48 (*Mus musculus*, Mammalia) eine entscheidende Rolle zu spielen (Zhao et al, 2007).

Interaktionsstudien mit dem PPC-lokalisierten sPub und den beiden sCdc48-Proteinen in *P. tricornutum* ergaben, dass sPub mit sCdc48-1 im PPC zu interagieren scheint (3.1.1.1). Betrachtet man den C-Terminus der Cdc48-Proteine aus *P. tricornutum*, fällt auf, dass sich das konservierte C-terminale DDDLYS-Motiv aus Mammalia auch bei hCdc48 und sCdc48-1 in leicht abgewandelter Form vorhanden ist (Abb. 4.1). Bei sCdc48-2 scheint dieses jedoch nicht konserviert zu sein, was die nicht nachweisbare Interaktion von sCdc48-2 mit sPub über Split-GFP erklären könnte.

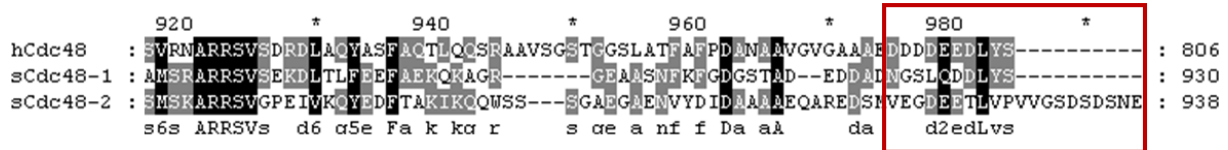


Abb. 4.1 Primärsequenz-Alignment des C-Terminus von Cdc48-Proteinen aus *P. tricornutum*. Das bei Mammalia konservierte DDDLYS-Motiv, welches höchstwahrscheinlich die Interaktion mit PUB-Domänen vermittelt (Zhao et al, 2007), findet sich in modifizierter Form auch bei hCdc48 und sCdc48-1 aus *P. tricornutum*. Bei sCdc48-2 ist dieses Motiv jedoch nicht konserviert (roter Rahmen).

Basierend auf der Lokalisation von sPub und dem Nachweis einer Interaktion des Proteins mit sCdc48-1 im PPC der Diatomee, kann für sPub eine Funktion innerhalb SELMA postuliert werden. Dabei wäre es möglich, dass sPub, aufgrund seiner Thioredoxin-Domäne, eine Redox-regulatorische Funktion im Rahmen des Importprozesses von Proteinen in das PPC an SELMA ausübt (Moog et al, 2011). Außerdem wäre denkbar, dass sPub in Chaperon-ähnlicher Weise, während dem Import von PPC-Proteinen über sCdc48-1/2, Disulfidbrücken an SELMA Substraten ausbildet.

Neben den bereits beschriebenen Komponenten konnten in einer aktuellen Studie zusätzliche putative symbiontische ERAD bzw. SELMA-Faktoren im PPC von *P. tricornutum* identifiziert und lokalisiert werden (Stork et al, 2012). Zu diesen zählen sNpl4, ein weiterer putativer Cofaktor von sCdc48(-1/-2) und sUfd1; sUbx, eine putative Andockstelle für sCdc48(-1/-2) an der PPM; sUbq, ein potentielles Homolog zu dem aus dem ERAD-System bekannten Ubiquitin-bindenden Enzyms Dsk2; sowie sPng1, eine putative de-glycosylierende Transglutaminase (Stork et al, 2012).

Die Faltung von importierten Proteinen im PPC durch Chaperone stellt eine wichtige Voraussetzung für die Erhaltung und Integrität des PPC-Proteoms dar. Außer an Proteinfaltungsprozessen sind Chaperone auch an Transportprozessen über Membranen bzw. Translokatoren in faltungsverhindernder Funktion, sowie an der Prävention von nicht-nativen Proteinaggregaten z.B. unter Stress (Hitzeschock, *heat shock*) beteiligt (Dierks et al, 1993; Hartl & Hayer-Hartl, 2002; Soll & Schleiff, 2004; Young et al, 2003). Mit sHsp70 konnte bereits ein PPC-residentes Chaperon identifiziert werden, welches diese Funktion wahrnehmen könnte (Gould et al, 2006b). Im Rahmen dieser Arbeit konnten drei weitere Chaperon-ähnliche Faktoren (sDtc, sDpc, sDnaK) im PPC der Diatomee identifiziert werden, die dort möglicherweise ebenfalls eine Rolle beim Transport und der Faltung von Proteinen

spielen (3.1.2, Abb. 4.2). Zwei dieser Proteine besitzen eine DnaJ-Domäne (sDtc und sDpc), die als Hsp70-Interaktionsdomäne gilt. Es ist bekannt, dass DnaJ-Domänen die ATPase-Funktion von Hsp70 stimulieren und somit auf den Faltungsprozess der Hsp70-Substrate Einfluss nehmen (Cyr et al, 1994; Laufen et al, 1999; Qiu et al, 2006). Es kann daher postuliert werden, dass es sich bei sDtc und sDpc aufgrund der Präsenz von DnaJ-Domänen möglicherweise um PPC-residente sHsp70-Cochaperone handelt.

sDpc besitzt zudem eine potentielle Protein-Disulfid-Isomerase-Domäne (PDI) welche sich zumeist bei ER-lokalisierten Chaperonen bzw. Protein-Disulfid-Isomerasen findet, die dort an der Ausbildung von Disulfidbrücken mitwirken (Wilkinson & Gilbert, 2004). Es kann daher angenommen werden, dass dieses Protein außerdem eine Rolle bei der Generierung von inter- bzw. intramolekularer Disulfidbrücken im Zuge der Faltung von PPC-Proteinen spielt. sDtc besitzt zusätzlich zur DnaJ-Domäne drei TPR-Motive (*tetratricopeptide repeats*) (3.1.2). Diese Motive vermitteln für gewöhnlich Protein-Protein-Interaktionen und finden sich u.a. bei Chaperonen, Zellzyklus- und Transkriptionsfaktoren, sowie bei Komponenten von Proteintransportkomplexen (Blatch & Lasse, 1999). Hinsichtlich Letzterem sind TPR-Motive z.B. bei den Importrezeptoren der ER-, plastidären, mitochondriellen und peroxisomalen Proteinimportmaschinerie von bedeutender Funktion (Schlegel et al, 2007). Da es bisher nicht gelang basierend auf dem TOC-vermittelten Transportmodell (Bullmann et al, 2010) einen Rezeptor für plastidäre Präproteine im PPC von *P. tricornutum* zu identifizieren und obwohl Homologien zwischen sDtc und den genannten Importrezeptoren äußerst gering ausfallen, wäre eine mögliche Funktion von sDtc als Präproteinrezeptor denkbar. sDtc könnte dabei zusammen mit sHsp70 bei der Erkennung und Weiterleitung plastidärer Präproteine zum putativen Translokator in der äußeren Membran der primären Plastide, ptOmp85, wirken (Moog, 2009). Um diese Hypothese zu prüfen könnten z.B. Interaktionsstudien zwischen sDtc und ptOmp85 Aufschluss geben.

Bei sDnaK handelt es sich um ein Chaperon des bakteriellen Hsp70-Typs, welches möglicherweise über horizontalen Gentransfer von einem Gammaproteobakterium akquiriert wurde. Homologe zu DnaK finden sich, mit Ausnahme verwandter Heterokontophyten (3.2), ausschließlich in Bakterien (BlastP). Da für sDnaK über BlastP lediglich eine Nukleotidbindedomäne und keine typische Substratbindedomäne auf Sequenzebene detektiert werden konnte (3.1.2), ist eine Rolle des Proteins als funktionelles Chaperon jedoch fraglich. Es kann dennoch nicht ausgeschlossen werden, dass DnaK eine mögliche Funktion bei der Faltung PPC-residenter Proteine wahrnimmt, oder etwa potentiell am Transport plastidärer

Proteine durch das PPC beteiligt ist. Hinsichtlich Letzterem ist, ausgehend vom SELMA- und TOC-basierten Translokationsmodell, die Erhaltung der Transportkompetenz plastidärer Proteine im PPC höchstwahrscheinlich eine wichtige Voraussetzung für die Überquerung der äußeren Membran der ehemals primären Plastide über ptOmp85, da der putative Translokator vermutlich nur von ungefalteten Proteinen passiert werden kann (Bullmann et al, 2010).

4.1.2 Besitzt das PPC eine proteolytische Aktivität?

Wie unter 3.1.3 dargestellt, gelang es im Rahmen dieser Arbeit ein unvollständiges eukaryotes Proteasom im PPC von *P. tricornutum* zu identifizieren. Dabei handelt es sich um Komponenten des 20S-Proteasomkomplexes von dem drei Alpha- (α 3-1, α 7-1 und α 7-2) und drei Beta-Untereinheiten (β 2, β 6 und β 7) als PPC-lokalisiert verifiziert werden konnten. Des Weiteren wurden bioinformatisch zwei zusätzliche Alpha-Untereinheiten (α 2, α 3-2) und eine Beta-Untereinheit (β 3) identifiziert (8.1). In einer aktuellen Studie gelang es außerdem das Set an 20S-Komponenten im PPC um eine weitere symbiontische Alpha-Untereinheit (*in silico*), sowie eine zusätzliche Beta-Untereinheit (β 1, PPC Lokalisation verifiziert) zu ergänzen (Stork et al, 2012). Insgesamt beläuft sich das putative 20S-Proteasom im PPC von *P. tricornutum* somit auf sechs Alpha- und fünf Beta-Untereinheiten (siehe Abb. 4.2) von denen einige der Beta-Untereinheiten (möglicherweise β 1, β 2) proteolytische Aktivität besitzen könnten (Baumeister et al, 1998; Groll et al, 1999).

Im Gegensatz zum nahezu vollständig detektierten Proteasom der Wirtszelle konnte in *P. tricornutum* kein regulatorischer 19S-Komplex im PPC identifiziert werden, der für gewöhnlich für die Substraterkennung, -deubiquitinierung und -weiterleitung in den 20S-Komplex verantwortlich ist (Voges et al, 1999). Außerdem ist unklar, ob sich die unvollständigen 20S-Komponenten strukturell assemblieren können um eine 20S-Proteasomtypische Fassstruktur im PPC auszubilden, weshalb eine proteolytische Aktivität der Faktoren fraglich ist. In Archaeen besteht das 20S-Proteasom nur aus zwei unterschiedlichen Untereinheiten (α und β), die sich in homoheptameren Ringen (Fassstruktur), ähnlich dem eukaryoten 20S-Proteasom (jeweils sieben unterschiedliche α - und β -Untereinheiten), formieren (Lowe et al, 1995). Es wäre daher möglich, dass fehlende Alpha- oder Beta-Untereinheiten im PPC von *P. tricornutum* durch andere, ähnliche Untereinheiten ersetzt werden um dort letztendlich als funktionelle Einheit zu assemblieren.

Es ist bekannt, dass das eukaryote 20S-Proteasom auch eigenständig, ohne 19S-Komplex, gewisse Proteine in einem Ubiquitin-unabhängigen Prozess degradieren kann (Baugh et al, 2009; Grune et al, 2003; Hoyt & Coffino, 2004; Orłowski & Wilk, 2003). Falls es zur Ausbildung eines 20S-Proteasomkomplex im PPC kommt, wäre eine proteolytische Aktivität somit auch ohne 19S regulatorischen Partikel (RP) denkbar. Allerdings stellt sich die Frage wie Substrate ohne 19S-RP entfalten werden, da der 20S-Komplex für gewöhnlich keine Kapazität zur Substratentfaltung aufweist (Baugh et al, 2009). Die Basis des 19S-RP wird bei Eukaryoten von sechs AAA-ATPasen (Rpt1-6) gebildet die beim 26S-Proteasom diese Aufgabe der Substratentfaltung übernehmen (Tomko & Hochstrasser, 2011). Im Vergleich dazu wird das 20S-Proteasom in Archaeen von homohexameren AAA-ATPasen des PAN-Typs, Homologen zu eukaryoten Rpt-ATPasen, reguliert (Medalia et al, 2009). Erst kürzlich konnte in Archaeen gezeigt werden, dass jedoch auch Cdc48 als Homohexamer mit dem 20S-Proteasomkomplex interagiert und ein funktionelles Proteasom ausbildet (Barthelme & Sauer, 2012). Die Autoren dieser Studie gehen hinsichtlich dieser Interaktion von einer ancestralen Funktion der Cdc48-Proteinfamilie aus (Barthelme & Sauer, 2012). Da im PPC von *P. tricornutum* zwei symbiontische sCdc48-Proteine lokalisieren und dort offensichtlich kein 19S-RP vorhanden ist (3.1.3), könnte möglicherweise eines der sCdc48-Proteine die Funktion des RP übernehmen und ein basales funktionelles Proteasom im PPC ausbilden (Moog et al, 2011). Dabei könnte sCdc48, ähnlich den Rpt-Komponenten aus Eukaryoten, bzw. den PAN-Faktoren aus Archaeen, proteasomale Substrate entfalten und dem putativen 20S-Komplex im PPC zur Degradation zuführen. Die Befähigung von Cdc48-Proteinen, Substrate in einer Chaperon-ähnlichen Weise zu entfalten, konnte bereits gezeigt werden (Beskow et al, 2009). Interaktionsstudien zwischen sCdc48-1/2 und den PPC-lokalisierten 20S-Komponenten könnten diesbezüglich weitere Hinweise liefern.

Das eukaryote 26S-Proteasom ist eng assoziiert an den Mechanismus der ER-assoziierten Degradation (ERAD). Im Zuge dieses Prozesses werden defekte Proteine im ER erkannt, Ubiquitin-abhängig vom Cdc48-Komplex (Cdc48/Ufd1/Npl4) ins Cytosol extrahiert, polyubiquitiniert und dem 26S-Proteasom zur Degradation zugeführt (Meusser et al, 2005). Da das SELMA-System auf die ursprüngliche ERAD-Maschinerie des Rotalgenendosymbionten zurückgeht (Felsner et al, 2011; Sommer et al, 2007), stellt sich die Frage nach einer funktionellen Verbindung von SELMA und den unvollständigen 20S-Komponenten im PPC von *P. tricornutum* (vgl. ERAD-26S-Proteasom). Basierend auf der aktuellen Datenlage ist eine solche Verbindung jedoch eher unwahrscheinlich: Zum einen konnten Homologe zu den ERAD-Komponenten Ufd2 und Rad23, welche für Substrat-Polyubiquitinierung bzw.

Substrat-Weiterleitung zum 19S-RP (in *P. tricornutum* nicht im PPC detektiert) benötigt werden (Farmer et al, 2010; Meusser et al, 2005), nur in der Wirtsversion und nicht als symbiontische Proteine in *P. tricornutum* identifiziert werden (Sommer et al, 2007; Stork et al, 2012). Zum anderen ist bei dem symbiontischen Ubiquitin (sUb) von *P. tricornutum* die Aminosäure Lys48, welche als bevorzugte Position für Polyubiquitinierung durch Ufd2 im Vorlauf proteasomaler Degradation gilt, nicht vorhanden (Sommer et al, 2007), was auch in anderen Linien der Heterokontophyten beobachtet werden konnte (Stork et al, 2012) und gegen eine Polyubiquitinierung im PPC spricht. Die Annahme, dass die Verwendung des symbiontischen Ubiquitin auf Proteinimportprozesse im Rahmen von SELMA beschränkt ist, wird durch die Präsenz von ptDUP im PPC von *P. tricornutum* gestützt, von dem postuliert wird, dass es SELMA-Substrate nach dem potentiell Ubiquitin-vermittelten Import ins PPC deubiquitiniert (Hempel et al, 2010).

In der Cryptophyte *Guillardia theta* findet sich, entgegen allen anderen bereits untersuchten Linien mit komplexen Plastiden rhodophytischen Ursprungs, ein nahezu vollständiges 26S-Proteasom im PPC (Douglas et al, 2001; Stork et al, 2012). Hinsichtlich der Tatsache, dass Cryptophyten einen Nucleomorph erhalten und somit ein PPC besitzen in dem Proteinbiosynthese stattfindet, ist das 26S-Proteasom dort wahrscheinlich in Mechanismen der Qualitätskontrolle im Kontext der Proteinexpression involviert (Douglas et al, 2001). Allerdings besitzen nicht alle Cryptophyten ein Proteasom im PPC. Genomische Analysen der Cryptophyte *Chroomonas mesostigmatica* CCMP1168 ergaben, dass ein PPC-lokalisiertes Proteasom in dieser Spezies zu fehlen scheint, während SELMA-Komponenten jedoch detektiert werden konnten (Moore et al, 2012). Des Weiteren finden sich ebenfalls keine proteasomalen Komponenten in den untersuchten Linien der Haptophyta und Apicomplexa im PPC (3.2), die dort jedoch über ein SELMA-System verfügen (Agrawal et al, 2009; Felsner et al, 2011; Spork et al, 2009; Stork et al, 2012). Es lässt sich daher vermuten, dass generell keine direkte bzw. notwendige Verbindung zwischen dem Import von Proteinen ins PPC über SELMA und den 20S-Proteasomkomponenten besteht (Stork et al, 2012).

Vielmehr könnte das PPC-residente Proteasom, falls proteolytisch aktiv, eine Rolle bei der Degradation fehlerhafter Proteine oder der Proteinprozessierung im PPC spielen. Allerdings liegen bisher in *P. tricornutum* (und anderen Heterokontophyten) keine Hinweise auf Mechanismen einer potentiellen Substraterkennung durch das PPC-lokalisierte putative Proteasom vor. Ob sCdc48-1/2 in diesem Kontext eine Rolle spielt und ob das symbiontische Ubiquitin in mögliche proteasomale Prozesse im PPC involviert ist, bleiben offene Fragen.

Letzteres ist jedoch aufgrund der zuvor geschilderten Datenlage eher unwahrscheinlich und lässt einen potentiell Ubiquitin-unabhängigen Degradationsmechanismus vermuten. Letztendlich bedarf es einer Reihe weiterer Studien um eine potentielle proteasomale Aktivität im PPC von *P. tricornutum* zu analysieren. Es kann zudem nicht ausgeschlossen werden, dass es sich bei den 20S-Komponenten im PPC möglicherweise um evolutionäre Relikte handelt.

Weitere Proteine, die eine mögliche proteolytische Aktivität im PPC von *P. tricornutum* ausüben könnten sind die unter 3.1.3 beschriebenen Trypsin-ähnliche Serinproteasen (sTlp1-6). Primäresequenzanalysen ergaben, dass es sich bei zwei der sTlps (sTlp2 und sTlp3) wahrscheinlich nicht um lösliche, sondern möglicherweise um Membranproteine mit einer C-terminalen Transmembrandomäne handelt. Diese sind, basierend auf ihren PPC-spezifischen BTS-Sequenzen (Tab. 8.2), somit höchstwahrscheinlich in der PPM verankert (Abb. 4.2).

Serin-Proteasen spielen eine Rolle bei einer Vielzahl zellulärer Prozesse. Sie wirken u.a. als digestive Enzyme (z.B. Trypsin), sowie als Regulatoren durch proteolytische Spaltung von Peptidbindungen im Zuge der Prozessierung von Zielproteinen (Kraut, 1977; Perona & Craik, 1995). Die Substratspezifität und die sich daraus ergebende Klassifizierung der verschiedenen Serin-Proteasen ist, neben strukturellen Determinanten, abhängig von der Beschaffenheit ihrer aktiven Zentren. Während bei Trypsin- bzw. Chymotrypsin-ähnlichen Serin-Proteasen die Aminosäuren His57, Asp102 und Ser195 konserviert sind und eine entscheidende Rolle für den Katalysemechanismus spielen (katalytische Triade), determinieren die AS an Position 189 (Asp bei Trypsin, Ser oder andere kleinere AS bei z.B Chymotrypsin), 216 und 226 (primäre Determinanten), sowie 190 und 228 (sekundäre Determinanten) die Substratspezifität (Perona & Craik, 1995). Trypsin-ähnliche Serin-Proteasen spalten für gewöhnlich Peptidbindungen die auf die AS Arg oder Lys folgen, während Chymotrypsin-ähnliche Serin-Proteasen Peptidbindungen nach aromatischen und z.T. hydrophoben AS hydrolysieren (Kraut, 1977).

Eine Analyse der Aminosäurekomposition der putativ aktiven Zentren der sTlp-Proteine ergab, dass es sich bei sTlp2, sTlp3, sTlp4 und sTlp5 (PPC-Lokalisation von sTlp5 nicht verifiziert) hinsichtlich ihrer Substratspezifität wahrscheinlich um Trypsin-ähnliche Proteasen handelt, während sTlp1 und sTlp6 eher als Chymotrypsin-ähnlich eingeordnet werden konnten (3.1.3; Abb. 3.9). Die Präsenz dieser putativen Proteasen im PPC deutet zunehmend auf eine proteolytische Aktivität in diesem Kompartiment hin. In jedem Fall setzt die PPC-Lokalisation der sTlps jedoch Regulationsmechanismen voraus, die diese Aktivität kontrollieren. Diesbezüglich wäre sowohl eine Regulation auf transkriptioneller Ebene, als

auch Proteinebene, z.B. über postrtranslationale Modifikationen oder auch proteinogene enzymatische Regulatoren, denkbar.

Zwar findet im PPC von *P. tricornutum* im Gegensatz zu Cryptophyten keine Proteinbiosynthese statt, nach dem Translokator-basierten Importmodell (Bullmann et al, 2010; Sommer et al, 2007) wird dieses jedoch wahrscheinlich von tausenden plastidären Präproteinen durchquert, wobei Mechanismen der Qualitätskontrolle von großem Vorteil sein können. Es wäre daher denkbar, dass die PPC lokalisierten sTlps, sowie die proteasomalen 20S-Komponenten eine Funktion dabei übernehmen. Des Weiteren fehlen bisher jegliche Anzeichen für PPC-lokalisierte Präsequenz-Proteasen, die womöglich TPL-Sequenzen nach erfolgtem Import von PPC-residenten Präproteinen ins PPC abspalten und degradieren. Auch diesbezüglich wäre eine Beteiligung der PPC-lokalisierten putativen Proteasen oder 20S-Komponenten vorstellbar. Weitere Studien sind allerdings notwendig um die potentielle Aktivität und Wirkungsweise der PPC-Lokalisierten Proteasen und proteasomalen 20S-Komponenten aufzuklären.

4.1.3 Das PPC beherbergt keine typischen Vesikel- und Cytoskelettelemente

Neben dem SELMA- und TOC-vermittelten Transportmodellen (Bullmann et al, 2010; Sommer et al, 2007) wurde in sekundär evolvierten Organismen hinsichtlich des Imports von Proteinen in die ehemals primäre Plastide ein vesikulärer Transportmechanismus durch das PPC postuliert (Gibbs, 1979). Dabei wird vermutet, dass Vesikel mit plastidären Proteinen von der PPM abknospen und über ein periplastidäres Retikulum im PPC zur äußeren Membran der primären Plastide (OEM) gelangen (Gibbs, 1979). Da sich diese Hypothese auf elektronenmikroskopische Beobachtungen stützte, wurde im Rahmen dieser Arbeit eine Analyse des postulierten vesikulären Transportmechanismus auf genomischer Ebene durchgeführt.

Wie in Tabelle 8.1 dargestellt, konnten über das genomische Screening jedoch entweder keine typischen Faktoren für vesikuläre, eukaryote Transportsysteme im PPC detektiert werden, oder *in vivo* lokalisierte Kandidaten zeigten eine nicht-PPC-Lokalisation. Weder Homologe zu klassischen eukaryoten COPI/II-, Clathrin-, noch ESCRT-Systemen konnten in einer symbiontischen Version (PPC) in *P. tricornutum* aufgefunden werden, während Wirtsfaktoren zumeist präsent waren (Tab. 8.1). Außer typischen eukaryoten Vesikelfaktoren konnten auch

keine PPC-spezifischen SNAREs, oder GTPasen (Rabs, GEFs, GAPs, Dynamin), die mit klassischen vesikulären Prozessen in Verbindung gebracht werden können, detektiert werden. Zwar konnte mit sDrp ein PPC-lokalisiertes Dynamin-ähnliches Protein identifiziert werden (3.1.4), es kann jedoch vermutet werden, dass dieses, aufgrund starker Homologie zu DRP5B-Proteinen des Teilungsapparates primärer Plastiden, dort eher in Prozesse der Plastidenteilung als in die Bildung von Transportvesikeln involviert ist (siehe 4.1.4). Dennoch kann eine Beteiligung von sDrp an putativen vesikulären Transportprozessen im PPC nicht ausgeschlossen werden. Generell sprechen die hier gesammelten Daten jedoch gegen einen vesikulären Transportmechanismus im PPC, der auf typischen eukaryoten Vesikelfaktoren basiert.

Neben den frühen elektronenmikroskopischen Beobachtungen vesikulärer Strukturen im PPC sekundär evolvierter Organismen mit komplexen Plastiden (Gibbs, 1979) existieren auch experimentelle Hinweise, die zumindest potentiell auf einen alternativen Transportweg für plastidäre Proteine in *P. tricornutum* hindeuten (Hempel et al, 2009). Dabei konnte in Studien mit dem Split-GFP-System gezeigt werden, dass der putative SELMA-Translokationskanal in der PPM der komplexen Plastide von *P. tricornutum*, bestehend aus sDer1-1/1-2, mit den Zielsteuerungssequenzen (BTS) von PPC-Proteinen interagiert, während dies nicht für die BTS-Sequenzen plastidärer (stromaler) Proteine zutrifft (Hempel et al, 2009). Diese Diskriminierung stromaler und PPC-residenter Proteine könnte auf eine alternative Transportroute (möglicherweise über Vesikel) für stromale Proteine hindeuten, könnte jedoch auch über eine sDer1-1/1-2-basierte Regulation beim Transportprozess beider Proteinklassen erklärt werden (Hempel et al, 2009).

Mögliche Komponenten, die in einen potentiell alternativen Transportweg im PPC involviert sein könnten, sind sORF261, sORF532a und sORF534 (siehe 3.1.7), Homologe zu nukleomorphkodierte ORFs der Cryptophyte *G. theta* mit genereller Konservierung in Organismen mit komplexen Plastiden rhodophytischer Herkunft und Rotalgen (siehe 3.2). Zwar konnten bioinformatisch für diese PPC-lokalisierten Proteine aus *P. tricornutum* keine funktionellen Domänen detektiert werden, Analysen mit den jeweiligen Homologen aus der Rotalge *C. merolae* (CMD101C, CMQ054C, CMI197C) lieferten jedoch Hinweise auf mögliche Funktionen der Proteine. So zeigte sich dabei, dass es sich bei CMD101C, CMQ054C und CMI197C um putative GTPasen mit Ähnlichkeit zu RHD3-Proteinen handelt (3.1.7). Die RHD3-Familie umfasst GTP-bindende Proteine (GTPasen), die weitläufig mit Proteinen der Dynamin-Familie verwandt sind und in Prozesse des Membrantransports und

der Membrandynamik involviert sind (Stefano et al, 2012). Bekannte Vertreter dieser Proteinfamilie sind z.B. RHD3 aus *Arabidopsis thaliana*, sowie Sey1p aus *Saccharomyces cerevisiae*, die eine Rolle bei der morphologischen Strukturierung des ERs spielen (Anwar et al, 2012; Stefano et al, 2012). Außerdem ist bekannt, dass diese Proteine nachweislich Einfluss auf Membran- bzw. Vesikelfusionsmechanismen ausüben (Anwar et al, 2012; Brands & Ho, 2002; Galway et al, 1997). Es wäre somit denkbar, dass die PPC-lokalisierten sORFs in *P. tricornutum* in diesem Kontext eine Rolle spielen. Allerdings ist die RHD3-Domäne der *C. merolae*-Pendants bei den drei PPC-lokalisierten sORFs (sORF261, sORF532a und sORF534) aus *P. tricornutum* (und *G. theta*) nicht detektierbar und somit weniger stark konserviert als andere proteininterne Sequenzbereiche (3.1.7), weshalb eine GTPase-Funktion der drei sORFs fraglich ist.

Hinweise auf eine generelle Involvierung des PPC-lokalisierten sORF534 in Transportmechanismen im PPC kommen jedoch aus aktuellen Studien in Apicomplexa. Diese besitzen ein Homolog zu sORF534, welches kürzlich im Zuge eines Screenings nach Apicoplast-lokalisierten Proteinen in *Toxoplasma gondii* identifiziert werden konnte, höchstwahrscheinlich im PPC lokalisiert ist und daher als PPP1 (*peripheral plastid protein 1*) bezeichnet wurde (Sheiner et al, 2011). Sheiner und Kollegen gelang es über eine konditionale Herunterregulierung der PPP1-Expression zu zeigen, dass es sich bei PPP1 in *T. gondii* um ein essentielles Protein handelt, welches eine entscheidende Rolle beim Transport nucleuskodierter Proteine in den Apicoplast (komplexe Plastide) spielt (Sheiner et al, 2011). Die Autoren kamen zu dem Schluss, dass PPP1 möglicherweise eine Funktion beim Proteinimport im Rahmen von SELMA einnimmt oder plastidäre Proteine (Apicoplast-Lumen) in Chaperon-ähnlicher Weise durch das PPC geleiten und dem putativen TOC-Komplex der OEM zuführen könnte (Sheiner et al, 2011). Da es sich bei sORF534 aus *P. tricornutum* um ein Homolog zu PPP1 aus *T. gondii* handelt, kann eine ähnliche Funktion und somit Beteiligung dieses Proteins an Transportprozessen im PPC postuliert werden. Ob sORF534 (und evtl. sORF261 und sORF532a) jedoch eine Funktion beim SELMA-basierten Proteintransport oder aber im Rahmen einer potentiell alternativen Transportroute, z.B über einen nicht konventionellen Vesikelmechanismus, im PPC wahrnimmt, bleiben offene Fragen, die es in zukünftigen Studien zu adressieren gilt.

Um die spezielle Morphologie des PPC zu gewährleisten (siehe 1.3), wäre die Präsenz von formgebenden bzw. formschaffenden Faktoren, d.h. Cytoskelett- oder Strukturelementen, im PPC zu erwarten. Darüber hinaus könnten mögliche PPC-lokalisierte Cytoskelett-

Komponenten im Kontext eines potentiellen Vesikeltransports im PPC, für den jedoch auf molekularer Ebene bisher keine Evidenz vorliegt, mitwirken. Allerdings gelang es im Rahmen dieser Arbeit nicht, typische Cytoskelettelemente im PPC von *P. tricornutum* nachzuweisen (3.1). Das genomische Screening resultierte, mit Ausnahme von Intermediärfilamenten, in der Detektion von Actin, α -, β - und γ -Tubulin, sowie einigen Myosinen (siehe auch (Heintzelman & Enriquez, 2010)), für die *in silico* cytosolische Lokalisationen vorhergesagt werden konnten (Tab. 8.1). Es handelt sich somit bei all diesen Proteinen höchstwahrscheinlich um Wirtsfaktoren. Kongruente Ergebnisse für Cytoskelett-Proteine im Cytosol von *P. tricornutum* wurden auch in einer anderen aktuellen Studie erzielt (Wickstead & Gull, 2011). Cryptophyten besitzen dahingegen Nucleomorph-kodierte Tubuline im PPC ihrer komplexen Plastide (Douglas et al, 2001; Keeling et al, 1999), die dort evtl. Cytoskelett-Funktionen wahrnehmen könnten. Aufgrund der besonderen Form des PPC von *P. tricornutum* (siehe Abb. 1.4) stellt sich jedoch die Frage, ob möglicherweise andere, nicht konventionelle Cytoskelettelemente im PPC existieren, die eine Rolle bei der Erhaltung der Morphologie der komplexen Plastide spielen könnten.

Potentielle Kandidaten könnten in diesem Kontext die PPC-lokalisierten Komponenten sPrp, sPel, sP4h und sSmc darstellen. Wie unter 3.1.8 beschrieben konnten diese Komponenten, entweder über Blast-Analysen oder die Detektion putativer funktioneller Domänen (NCBI-CDS), mit Proteinen anderer Organismen in Verbindung gebracht werden, die mögliche strukturelle Funktionen ausüben. Darunter Trans-Sialidasen (sPrp), die im Kontext von Zell-Zell-Interaktionen als Oberflächendeterminanten wirken (Frasch, 2000); Pectin(methyl)esterasen (sPel), welche eine Rolle bei der Adhäsion von pflanzlichen Zellen spielen (Micheli, 2001); Prolin 4-Hydroxylasen (sP4h), die in Mammalia durch Hydroxylierung von proteininternen Prolinresten die Collagen-Tripelhelix (strukturelle Matrix) stabilisieren (Gorres & Raines, 2010) und schließlich SMC-Proteine (sSmc), die Ähnlichkeit zu Motorproteinen wie Kinesin und Myosin aufweisen (Strunnikov & Jessberger, 1999) und die für gewöhnlich in Prozesse der Chromosomen-Segregation involviert sind (Harvey et al, 2002). Generell wäre eine Beteiligung von sPrp, sPel, sP4h und sSmc an strukturellen Funktionen im PPC von *P. tricornutum* möglich, diese Vermutung gilt jedoch im Weiteren experimentell untersucht zu werden. Es wäre zudem denkbar, dass die PPC-lokalisierten sORFs, alternativ zu den zuvor diskutierten möglichen Funktionen (siehe oben), eine Rolle bei strukturellen Prozessen im PPC der Diatomee spielen (evtl. Membranorganisation).

Neben den bisher genannten Faktoren, gelang es über das bioinformatische Screening einige Proteine mit putativen Ankyrin-Domänen und potentiellen PPC-spezifischen BTS-Sequenzen zu detektieren (Tab. 8.1). Ankyrine sind als Cytoskelettelemente bekannt, Ankyrin-Domänen (*repeats*) finden sich jedoch außer bei Cytoskelett-organisierenden Proteinen auch bei einer Reihe regulatorischer Komponenten, wo sie als Plattformen für molekulare Interaktionen gelten (Kordeli et al, 1995; Mosavi et al, 2004; Sedgwick & Smerdon, 1999). Lokalisationsstudien mit einem der PPC-Kandidaten ergaben, dass das Protein wahrscheinlich mitochondriell lokalisiert (Tab. 8.1). Es ist jedoch nicht auszuschließen, dass eines der übrigen putativen Ankyrine (Tab. 8.1) zu einer Cytoskelett-ähnlichen oder strukturellen Funktion im PPC von *P. tricornutum* beiträgt oder andere nicht typische Faktoren in diesem Kontext involviert sind. Ergänzende Studien könnten diesbezüglich Aufschluss geben.

4.1.4 Komplexe Plastide - simple Teilung?

Im Gegensatz zu primären Plastiden ist hinsichtlich molekularer Mechanismen und beteiligter Faktoren am Teilungsprozess komplexer Plastiden bisher nur wenig bekannt. In *P. tricornutum* findet die Teilung bzw. Segregation der komplexen Plastide direkt zu Beginn der Mitosephase vor Einleitung der Karyo- und Cytoskinese statt (Bowler et al, 2010). Zuerst teilt sich dabei die frühere primäre Plastide gefolgt von der Teilung der PPM und der cER-Membran (Hashimoto, 2005). Da die komplexe Plastide von *P. tricornutum* auf eine internalisierte Rotalge zurückgeht, kann vermutet werden, dass Faktoren des ursprünglichen Teilungsapparates des Endosymbionten erhalten wurden.

Im Rahmen dieser Arbeit wurde der Focus speziell auf die Analyse möglicher teilungsrelevanter Proteine im PPC der Diatomee gerichtet. Dabei gelang es mit sDrp und sPdr1 zwei Proteine mit Ähnlichkeiten zu cytosolischen Plastidenteilungsfaktoren aus Rotalgen im PPC der komplexen Plastide von *P. tricornutum* zu identifizieren (3.1.4.1). Bei sDrp handelt es sich um ein Dynamin-ähnliches Protein mit starker Homologie zu Proteinen der DRP5-Klasse, welche in primären Plastiden eine bedeutende Rolle beim Plastidenteilungsprozess (Abschnürung) auf cytosolischer Seite wahrnehmen (Gao et al, 2003; Miyagishima et al, 2003; Yoshida et al, 2006). sDrp besitzt eine starke Sequenzidentität mit Dnm2, dem DRP5B-Pendant aus *C. merolae* (3.1.4.1). Letzteres bildet eine ringartige Struktur um die Plastide der Rotalge während der Teilungsphase aus, welche sich im Zuge der Division kontrahiert und anschließend wieder auflöst (Miyagishima et al, 2003). Blast-

Analysen deuten auf einen rhodophytischen Ursprung von sDrp hin. Es ist daher anzunehmen, dass sDrp, ortholog zu DRP5B-Proteinen aus Organismen mit primären Plastiden, im PPC am Teilungsprozess der ehemals primären Plastide mitwirkt (Abb. 4.2) und einen endosymbiontischen Ursprung (Rotalge) aufweist. Anders verhält es sich z.B. in den Apicomplexa, bei denen die ursprüngliche Teilungsmaschinerie des Endosymbionten wahrscheinlich verloren ging und welche somit kein Homolog zu DRP5 besitzen. Dort teilt sich der Apicoplast durch Beteiligung eines höchstwahrscheinlich von Seiten des Wirts rekrutierten Proteins der Dynamin-Klasse (DrpA in *T. gondii*), welches phylogenetisch von der DRP5B/ARC5-Klasse unterscheidet (van Dooren et al, 2009). Erste Evidenzen für eine Involvierung von sDrp in Prozesse der Plastidenteilung im PPC von *P. tricornutum* konnten bereits durch Expression einer dominant-negativen Mutante von sDrp als eGFP-Fusionsprotein beobachtet werden. Dabei deutete sich in mikroskopischen Analysen an, dass die Überexpression einer sDrp-Variante mit Mutationen in der GTPase-Domäne den Plastidenteilungsprozess zu beeinträchtigen scheint (nicht gezeigt).

Zusätzlich zu DRP5B konnte in Rotalgen kürzlich ein weiteres cytosolisches Protein identifiziert werden, dass eine bedeutende Funktion bei der Teilung der primären Plastide einnimmt. Dabei handelt es um PDR1 (*plastid-dividing ring 1*), eine Glycogenin-ähnliche Glycosyltransferase der GT8-Familie, welche höchstwahrscheinlich den äußeren (cytosolischen) PD-Ring (*plastid-dividing*) konstruiert (Yoshida et al, 2010). PD-Ringe gelten als universelle Strukturen im Kontext plastidärer Division (präsent in primären und komplexen Plastiden), die Zusammensetzung ihrer Struktur war jedoch lange Zeit unbekannt. Aktuelle Studien in *C. merolae* ergaben, dass der PD-Ring der Rotalge aus Polyglucanfilamenten (Glucose) besteht, die von PDR1 auf cytosolischer Seite am Ort der Plastidenteilung assembliert werden (Yoshida et al, 2010). PDR1 konnte dabei nicht nur in konstruierender Weise, sondern auch als möglicher struktureller Bestandteil der PD-Maschinerie identifiziert werden (Yoshida et al, 2010).

Zwar konnten durch Blast-Analysen keine direkten Homologe zu PDR1 aus *C. merolae* in der Genomdatenbank von *P. tricornutum* identifiziert werden, es gelang allerdings über direkte Suche nach Glycogeninen fünf Proteine mit Glycogenin-Domäne zu detektieren. Wie unter 3.1.4.1 dargetellt, konnte bisher für eines dieser Proteine (sPdr1) über *in vivo* FL-GFP-Lokalisationsstudien eine PPC-Lokalisation nachgewiesen werden, die sich darüberhinaus auch über das Self-Assembling-GFP System bestätigen ließ (siehe Abb. 3.12). Aufgrund ähnlicher Domänenstruktur beider Proteine (siehe Abb. 3.11B) kann postuliert werden, dass

es sich bei sPdr1 möglicherweise um ein Ortholog zu PDR1 aus *C. merolae* handelt und dieses somit ebenso in Prozesse der Plastidenteilung involviert sein könnte (Abb. 4.2).

Glycogenine sind für gewöhnlich als Primer für die Glycogensynthese ausgehend von UDP-Glucose im Kontext der Energiespeicherung (z.B. in Mammalia) bekannt (Lomako et al, 2004). Diese zur Autoglycosylierung befähigten Enzyme binden 8-20 Glucosereste über eine O-glycosidische Bindung (Tyr194) zunächst an sich selbst und stellen die Polysaccharidkette im Weiteren der Glycogensynthese zur Verfügung (Issoglio et al, 2012; Lomako et al, 2004; Roach et al, 2012). In *P. tricornutum* findet die Speicherung von Energie jedoch in Form von Lipiden oder durch Synthese von Chrysolaminarin (β -1,3-Glucan), wahrscheinlich ebenfalls ausgehend von UDP-Glucose, in der Vakuole statt (Kroth et al, 2008). Es kann daher vermutet werden, dass das PPC-lokalisierte sPdr1 keine Funktion im Kontext der Energiespeicherung wahrnimmt, zumal auch keine weiteren Enzyme, welche sich mit der Synthese von Glycogen in Verbindung bringen ließen (z.B. Glycogensynthase), als PPC-spezifisch in *P. tricornutum* identifiziert werden konnten.

Um zu untersuchen ob sPdr1 ähnlich wie PDR1 aus *C. merolae* zur Autoglycosylierung befähigt ist, wurde eine Analyse des Glycosylierungsstatus von sPdr1 durchgeführt. Diese ergab, dass das in *P. tricornutum* exprimierte sPdr1-eGFP Fusionsprotein scheinbar glycosyliert vorliegt (3.1.4.2, Abb. 3.13). Allerdings handelt es sich bei der beobachteten Glycosylierung höchstwahrscheinlich eher um eine N- als eine für Glycogenine typische O-Glycosylierung, zumal diese nach Behandlung mit PNGase (spaltet spezifisch N-glycosidische Bindungen) im *Glyco-Stain* nicht mehr nachweisbar war (siehe Abb. 3.13). Gleiche Ergebnisse konnten für sPdr1-eGFP auch nach Inhibition der N-Glycosylierungsprozesse im ER durch Tunicamycin erzielt werden (Madeleine Peschke, unveröffentlicht). Für PDR1 aus *C. merolae* konnte gezeigt werden, dass dieses ausschließlich während der Plastidenteilungsphase glycosyliert vorliegt (Yoshida et al, 2010). Gleiches könnte auch für sPdr1 aus *P. tricornutum* zutreffen, was erklären würde warum für aufgereinigtes sPdr1-eGFP Fusionsprotein aus einer nicht synchronisierten Kultur keine O-Glycosylierung nachweisbar war (siehe Abb. 3.13). Es ist daher im Weiteren zu klären, ob bei Verwendung einer synchronisierten *P. tricornutum* Kultur während der Plastidenteilungsphase eine O-Glycosylierung bei sPdr1-eGFP detektiert werden kann. Die beobachtete potentielle N-Glycosylierung des PPC-lokalisierten sPdr1-eGFP Fusionsproteins deutet jedoch darauf hin, dass möglicherweise auch N-glycosylierte Proteine in das PPC der Diatomee importiert werden können (Madeleine Peschke, unveröffentlicht).

Ob es sich bei sPdr1 tatsächlich um ein Ortholog zu PDR1 aus *C. merolae* handelt, konnte in dieser Arbeit nicht eindeutig geklärt werden. Interaktionsstudien von sPdr1 und sDrp könnten Hinweise liefern ob sPdr1 Bestandteil der PPC-lokalisierten Plastidenteilungsmaschinerie ist. Es kann nicht ausgeschlossen werden, dass die putative Glycosyltransferase sPdr1 eine Funktion fernab von Plastidenteilungsprozessen im PPC der Diatomee wahrnimmt. Diese könnten bisher unbekannte metabolische oder strukturelle Funktionen im PPC miteinschließen, bei denen möglicherweise die Ausbildung von Polysacchariden (putative Glycogenin-Funktion) relevant ist.

Interessanterweise ergab eine proteomische Analyse der isolierten PD- (*plastid dividing*) Maschinerie aus *C. merolae*, dass neben PDR1, Dnm2, FtsZ und u.a. Toc75 auch CMD101C, das Homolog zu dem PPC-lokalisierten sORF261 aus *P. tricornutum* detektiert werden konnte. Wie bereits erwähnt (3.1.7 und 4.1.3), handelt es sich bei CMD101C um eine potentielle GTPase der RHD-Familie, die möglicherweise in Prozesse der Membranorganisation involviert ist. Aufgrund der koinzidenten Aufreinigung von CMD101C mit der PD-Maschinerie wäre es denkbar, dass diesem eine mögliche Rolle bei der Teilung der primären Plastide in *C. merolae* zukommt. Ob ähnliches jedoch auch für sORF261 in *P. tricornutum* zutrifft bleibt zu klären. Interaktionsstudien mit dem PPC-lokalisierten sDrp und sORF261 könnten Aufschluss über eine mögliche Beteiligung des Proteins an Teilungsprozessen der ehemals primären Plastide im PPC geben.

Hinsichtlich weiterer potentieller Faktoren des Teilungsapparates der komplexen Plastide von *P. tricornutum*, gelang es, neben sDrp und sPdr1, drei putativ stromale FtsZ-Proteine *in silico* zu identifizieren (Tab. 8.1). FtsZ-GTPasen spielen eine wichtige Rolle beim Teilungsprozess primärer Plastiden, bei dem sie als zentrale Komponenten des stromalen Teilungsapparates filamentöse Ringstrukturen an der stromalen Seite der IEM ausbilden (Yoshida et al, 2012). Ähnlich wie bei Rotalgen (*C. merolae*) (Miyagishima & Kabeya, 2010; Yang et al, 2008) finden sich in *P. tricornutum* keine Homologe zu anderen an der Teilung primärer Plastiden (Grünalgen und Pflanzen) beteiligter Faktoren wie Komponenten des plastidären Min-Systems, ARC3, ARC6, PARC6, MCD1, PDV1, PDV2 oder GC1. Es kann daher angenommen werden, dass die drei FtsZ-Proteine Hauptbestandteile des stromalen Plastidenteilungsapparates in *P. tricornutum* darstellen. Aufgrund der Präsenz von sDrp und sPdr1 im PPC von *P. tricornutum* kann postuliert werden, dass die Kernkomponenten der Plastidenteilungsmaschinerie des rhodophytischen Endosymbionten weitestgehend beibehalten wurden. Ob weitere Faktoren involviert sind oder ob es sich beim Teilungsapparat

komplexer Plastiden im Gegensatz zu dem primärer Plastiden von Grünalgen und Pflanzen um eine bislang „relativ simple“ Maschinerie handelt, werden zukünftige Studien zeigen.

4.1.5 Metabolische Prozesse im PPC sind begrenzt

Wie unter 1.3 beschrieben, sind bisher nur wenige Hinweise auf eine metabolische Relevanz des PPC von *P. tricornutum* (Präsenz metabolischer Enzyme im PPC) vorhanden. Jedoch konnte ein Enzym des oxidativen Pentosephosphatwegs (OPP), welcher in *P. tricornutum* hauptsächlich im Cytosol abläuft, als PPC-lokalisiert nachgewiesen werden (Gruber et al, 2009). Dabei handelt es sich um s6PGDH (6-Phosphogluconolacton-Dehydrogenase), ein Enzym, welches die Umwandlung von Gluconat-6-Phosphat unter Entstehung von NADPH und CO₂ zu Ribulose-5-Phosphat im Rahmen des OPP katalysiert (Gruber et al, 2009). Zumal Prozesse wie Glycolyse und Pentosephosphatweg in Bezug auf die subzelluläre Verteilung beteiligter Enzyme in *P. tricornutum* bereits bioinformatisch und z.T. auch über *in vivo* Lokalisationsstudien analysiert wurden (Gruber et al, 2009; Kroth et al, 2008), wurden diese in dieser Arbeit nicht behandelt. Vielmehr war der Fokus auf die Identifizierung möglicher PPC-residenter Carboanhydrasen, im Kontext der Konzentrierung anorganischen Kohlenstoffs (CCM, *carbon concentrating mechanism*), gerichtet. Kroth et al. (2008) berichteten u.a. über sieben Carboanhydrasen (CAs), die im Rahmen eines genomischen Screenings in *P. tricornutum* identifiziert (Kroth et al, 2008) und von denen bereits zwei β -CAs (PtCa1 und PtCA2) im Plastidenstroma der Diatomee lokalisiert werden konnten (Kitao et al, 2008; Tanaka et al, 2005). Carboanhydrasen spielen u.a. eine wichtige Rolle beim Prozess der Photosynthese und dabei im speziellen bei der Fixierung von CO₂ in der Plastide (Badger & Price, 1994). Sie katalysieren die reversible Umwandlung von HCO₃⁻ zu CO₂ und gewährleisten dadurch eine effiziente Versorgung von RuBisCO (Ribulose-1,5-Bisphosphat-Carboxylase/-Oxygenase) mit anorganischem Kohlenstoff in Form von CO₂ (Kaplan & Reinhold, 1999). Es wird vermutet, dass CAs in Diatomeen eine bedeutende Rolle im Zuge eines CCM zukommt (Hopkinson et al, 2011; Kitao et al, 2008; Tanaka et al, 2005).

Im Rahmen dieser Arbeit gelang es mittels *in vivo* eGFP-Lokalisationsstudien zwei Alpha-Carboanhydrasen (α Ca1 und α Ca2) im PPC von *P. tricornutum* nachzuweisen (3.1.6, Abb. 4.2), während zwei weitere α CAs im ER der Diatomee lokalisiert werden konnten (Tab. 8.2). Diese Ergebnisse sind jedoch nicht vollständig mit denen einer anderen aktuellen Studie über die Lokalisation dieser Carboanhydrasen in *P. tricornutum* im Einklang (Tachibana et al,

2011). Während für α Ca1 in beiden Studien eine PPC-Lokalisation nachgewiesen werden konnte (gleiches gilt für die ER-Lokalisation von α Ca2), unterscheiden sich die Ergebnisse hinsichtlich der Lokalisationen von α Ca2 (hier PPC, dort ER) und α Ca1 (hier ER, dort PPC). Dies könnte möglicherweise mit den gewählten Längen der für *in vivo* GFP-Lokalisationsstudien eingesetzten Präsequenzen erklärt werden. Felsner et al. (2010) konnten in diesem Kontext zeigen, dass bei sehr kurzen TPL-Sequenzen plastidärer Proteine auch Teile des maturen Proteins zur Funktion der Transitpeptid-ähnlichen Sequenzen in *P. tricornutum* beitragen können (Felsner et al, 2010). Da die Längen der Präsequenzen der Carboanhydrasen nicht eindeutig bestimmt werden konnten, wurde in dieser Arbeit für die Expression von Präsequenz-GFP-Fusionsproteinen eine Länge von 169 AS (α Ca2) und 62 AS (α Ca1) verwendet, wobei von Tachibana et al. (2011) diesbezüglich 46 (α Ca2) und 45 (α Ca1) AS vor GFP exprimiert wurden. Die Expression einer „verkürzten“ BTS fusioniert mit GFP könnte sich dabei auf die Ziesteuerungsprozesse ausgewirkt haben und die abweichenden Resultate erklären. Die Expression von FL-eGFP-Fusionsproteinen von α Ca2 und α Ca1 könnte die tatsächliche Lokalisation der beiden Proteine höchstwahrscheinlich klären.

Trotz teilweise divergenter Ergebnisse, deutet die Präsenz putativer Carboanhydrasen im PPC von *P. tricornutum* auf eine metabolische Relevanz dieses Kompartiments hin. Durch die Architektur der komplexen Plastide (vier Hüllmembranen) sind Versorgungsmechanismen des plastidären Kohlenstofffixierungssystems (RuBisCO) im Vergleich zu primären Plastiden (zwei Hüllmembranen) denkbar erschwert. Das PPC als abgeschlossener Reaktionsraum, der die ehemals primäre Plastide direkt umgibt, könnte in diesem Fall als eine Art Barriere für entweichendes CO_2 aus der Plastide fungieren (Moog et al, 2011). Die PPC-residenten Carboanhydrasen (siehe auch Abb. 4.2) könnten in diesem Kontext die CO_2 -Konzentration in unmittelbarer Nähe der ehemals primären Plastide erhöhen und, ergänzend zu den stromalen Beta-Carboanhydrasen (Kitao et al, 2008; Tanaka et al, 2005), eine direkte Versorgung von RuBisCO mit anorganischem Kohlenstoff gewährleisten (Kroth et al, 2008). Neben möglichem aktiv oder passiv konzentriertem $\text{CO}_2/\text{HCO}_3^-$, ergibt sich durch die von s6PGDH katalysierte Reaktion eine weitere potentielle, für die α CAs relevante, CO_2 -Quelle im PPC (siehe 1.3). Generell deuten die hier und in anderen Studien (Gruber et al, 2009; Kroth et al, 2008) erzielten Ergebnisse darauf hin, dass metabolische Prozesse im PPC von *P. tricornutum* stark begrenzt sind. Anders verhält es sich z.B. in Cryptophyten (*G. theta*), wo die Biosynthese und Speicherung von Stärke im PPC abläuft (Gilson, 2001; Haferkamp et al, 2006).

4.1.6 Homöostase und Regulationsmechanismen im PPC

Neben einer Vielzahl von Proteinen benötigen Plastiden auch Phospholipide von Seiten der Wirtszelle. Die finden sich zumeist in Form von ER-synthetisiertem Phosphatidylcholin (PtdCho) auf der dem Cytosol zugewandten Seite der äußeren Membran primärer Plastiden, wo sie neben typischen plastidär-synthetisierten Glycolipiden existieren (Dorne et al, 1985). Es wird vermutet, dass dies die Integration des ursprünglichen Cyanobakteriums als Organell (Plastide) reflektiert und gilt als Notwendigkeit der Adaption einer Importmaschinerie für Proteine von Seiten der Wirtszelle (Cavalier-Smith, 2000; Maréchal et al, 1997). In diesem Kontext kann davon ausgegangen werden, dass bereits früh nach dem Ereignis der primären Endosymbiose ein Verlust cyanobakterieller Lipopolysaccharide und Lipoproteine zu verzeichnen war und diese, vermutlich über cytosolische Lipidtransferproteine (LTPs), durch Phospholipidspezies der Wirtszelle ersetzt wurden (Cavalier-Smith, 2000). Die Versorgung der äußeren Membran (OEM) primärer und komplexer Plastiden mit ER-synthetisierten Phospholipiden des Wirts ist somit höchstwahrscheinlich von kontrollierten Transfermechanismen abhängig. Grundsätzlich wären für einen solchen Transfer, neben dem aktiven Transport über Lipidtransferproteine, zwei weitere Mechanismen denkbar, nämlich der Transport über Vesikel, sowie ein Austausch von Phospholipiden über Membrankontaktstellen (Griac, 2007; Lev, 2010). Da bisher jedoch keine Faktoren identifiziert werden konnten, die auf einen typischen Vesikelverkehr im PPC hindeuten (3.1, Tab. 8.1), ist davon auszugehen, dass der Transfer von Phospholipiden zwischen ER und OEM eher über Membrankontaktstellen oder LTPs abläuft.

In dieser Arbeit gelang es mit sSec14 ein putatives Lipidtransferprotein im PPC der Diatomee von *P. tricornutum* zu identifizieren (siehe 3.1.5), welches dort möglicherweise in Lipidtransferprozesse involviert sein könnte (Abb. 4.2). sSec14 wurde nach seinem Homolog Sec14p aus *S. cerevisiae* benannt, welches *in vitro* dazu befähigt ist Phospholipide wie Phosphatidylinositol und Phosphatidylcholin zwischen artifiziellen biologischen Membranen zu transferrieren, während es *in vivo* an Sekretionsprozessen, sowie der Regulation von Lipidbiosynthese und -turnover beteiligt ist (Schnabl et al, 2003). Da keine Faktoren für die Biosynthese von Glycero(phospho)lipiden im PPC von *P. tricornutum* identifiziert werden konnten (siehe 3.1.5 und 8.1), ist eine unmittelbare Beteiligung von sSec14 an Regulationsmechanismen der Lipidsynthese, welche in *P. tricornutum* höchstwahrscheinlich am ER, in der ehemals primären Plastide und den Mitochondrien abläuft (Tab. 8.1), eher unwahrscheinlich. Die Präsenz des putativen Lipidtransferproteins sSec14 im PPC lässt

jedoch vermuten, dass dieses im Rahmen der Homöostase der ehemals primären Plastide eine Funktion bei der Versorgung der OEM mit ER-synthetisierten Phospholipiden (wahrscheinlich PtdCho) spielen könnte. Dabei könnte sSec14 über aktiven Lipidtransfer eine Verbindung zwischen PPM und OEM herstellen. Allerdings müsste, auf dieser Vermutung basierend, auch ein Lipidtransfer von der cER-Membran zur PPM erfolgen, da diese die cER-Membran von der OEM der ehemals primären Plastide abgrenzt. Möglicherweise wird diese Funktion von einem weiteren Sec14-ähnlichen Protein aus *P. tricornutum* übernommen, *in silico* Analysen deuten jedoch nicht auf die Lokalisation eines solchen Proteins im ER hin (siehe 3.1.5). Außerdem wären Faktoren in der PPM zu erwarten, welche in der Lage sind die Orientierung von Lipidspezies in der PPM-*bilayer* (Membrandoppelschicht) zu verändern um diese für Transportvorgänge zugänglich zu machen. Dies könnte möglicherweise über Flippasen gewährleistet werden (Daleke, 2007). Prinzipiell wäre auch eine Kombination von Lipidtransfer über Membrankontaktstellen (cER-Membran-PPM) und aktivem Transport, evtl. über sSec14 (PPM-OEM), denkbar. Ob sSec14 jedoch tatsächlich in der Lage ist aktiv Phospholipide zu transferieren und eine Rolle bei der Homöostase der OEM der komplexen Plastide von *P. tricornutum* spielt, muss in zukünftigen Studien geklärt werden.

Mechanismen der Regulation und Signaltransduktion sind wichtige Prozesse um auf äußere und innere Reize zu reagieren und z.B. die Expression von Genen, oder die Aktivität von Proteinen eines Cytoplasmas bzw. einer Zelle zu steuern. Solche Regulationsmechanismen können auf unterschiedlichste Arten vollzogen werden und können dabei z.T. entscheidende Einflüsse auf räumliche und zeitliche Prozesse, wie z.B. den circadianen Rhythmus, ausüben (Coesel et al, 2009). Bekannte Formen von Signaltransduktion und Regulation stellen z.B. die Phosphorylierung von Proteinen über Kinasen, Photorezeption (z.B. über Cryptochrome) und Redox-regulatorische Prozesse (Thioredoxin- und Glutathion-Systeme) dar (Cashmore et al, 1999; Grant, 2001; Rubin & Rosen, 1975). Im Rahmen dieser Arbeit konnten jedoch in *P. tricornutum* keine neuen Faktoren in den genannten Kategorien als PPC-lokalisiert identifiziert werden (3.1 und Tab. 8.1). Wie in Tabelle 8.1 dargestellt, gelang es in den meisten Fällen lediglich Proteine mit cytosolischer Lokalisationsvorhersage, womöglich Wirtsfaktoren, zu detektieren. Im Hinblick auf die Präsenz potentieller Redox-regulatorischer Komponenten im PPC von *P. tricornutum* konnten bereits in früheren Studien Thioredoxin (sTrxH) und eine NADPH-abhängige Thioredoxin-Reduktase identifiziert werden, bezüglich möglicher Zielproteine liegen jedoch bisher keine Hinweise vor (Weber et al, 2009). Zudem wurde in dieser Arbeit mit dem sCdc48-1-Interaktionspartner sPub ein weiteres PPC-Protein mit Thioredoxin-Domäne detektiert (3.1.1), für das eine mögliche Beteiligung an

regulatorischen Prozessen nicht ausgeschlossen werden kann. Basierend auf der aktuellen Datenlage kann daher davon ausgegangen werden, dass Kinasen-vermittelte Regulations- oder Signalprozesse, sowie Photorezeption keine prominente Rolle im PPC von *P. tricornutum* spielen, während eine Thiothionin-basierte Redox-Regulation offenbar stattzufinden scheint.

4.2 Das PPC der Diatomee *Phaeodactylum tricornutum*

Die Kenntnisse um das PPC-Proteom der Diatomee *P. tricornutum* konnte im Rahmen dieser Arbeit, ausgehend von 11 Proteinen, um weitere 29 durch *in vivo* Lokalisation verifizierte Proteine erweitert werden (Abb. 4.2). Zusammen mit einer weiteren aktuellen Studie (Stork et al, 2012) konnten somit bisher insgesamt 45 Proteine im PPC der Diatomee lokalisiert werden, wobei *in silico* Hinweise zu mindestens 8 weiteren Faktoren vorliegen. Die generierten Daten ermöglichen weitere Einblicke in bereits bekannte Prozesse, wie SELMA-vermittelten Proteinimport, Proteinfaltung und Metabolismus im PPC. Des Weiteren konnten jedoch auch neue potentielle Funktionen hinsichtlich Proteolyse und Degradation, Plastidenteilung und Lipidtransfer im PPC von *P. tricornutum* identifiziert werden (Abb. 4.2). Gleichbedeutend wichtig sind Prozesse und Faktoren die nicht im PPC detektiert werden konnten. Dazu zählen u.a. typische eukaryote Vesikel- und Cytoskelettelemente und klassische Signaltransduktionsfaktoren.

Da das PPC ein auf natürliche Weise minimiertes ehemaliges Cytoplasma eines Rotalgenendosymbionten darstellt (Cavalier-Smith, 2002), kann nach bisherigem Kenntnisstand davon ausgegangen werden, dass der Großteil klassischer cytosolischer Funktionen und metabolischer Abläufe drastisch reduziert wurde bzw. im Zuge der Koevolution von Wirt und Symbiont verloren ging. Dies ist jedoch hinsichtlich der Tatsache, dass es sich beim PPC um ein nicht autonomes Cytoplasma handelt, das von der Zufuhr nucleuskodierter Proteine der Wirtszelle abhängig ist, nicht erstaunlich. So kann davon ausgegangen werden, dass redundante Prozesse und Faktoren weitestgehend eliminiert worden sind. Spuren typisch cytosolischer Prozesse (Proteintransport, -faltung und degradation) sind zwar vorhanden, diese scheinen jedoch spezifisch an Prozesse angepasst worden zu sein, die sich aus der Notwendigkeit der stabilen Integration des Endosymbionten in die Wirtszelle ergaben. Ein Beispiel für spezifische funktionelle Adaption stellt dabei das SELMA-System dar, welches ursprünglich vermutlich Proteindegradationsprozesse im

Rahmen von ERAD am ER des Endosymbionten vollzog und nun, höchstwahrscheinlich degradationsunabhängig, Proteinimportfunktionen an der PPM ausübt (Bolte et al, 2011; Hempel et al, 2009; Sommer et al, 2007; Stork et al, 2012).

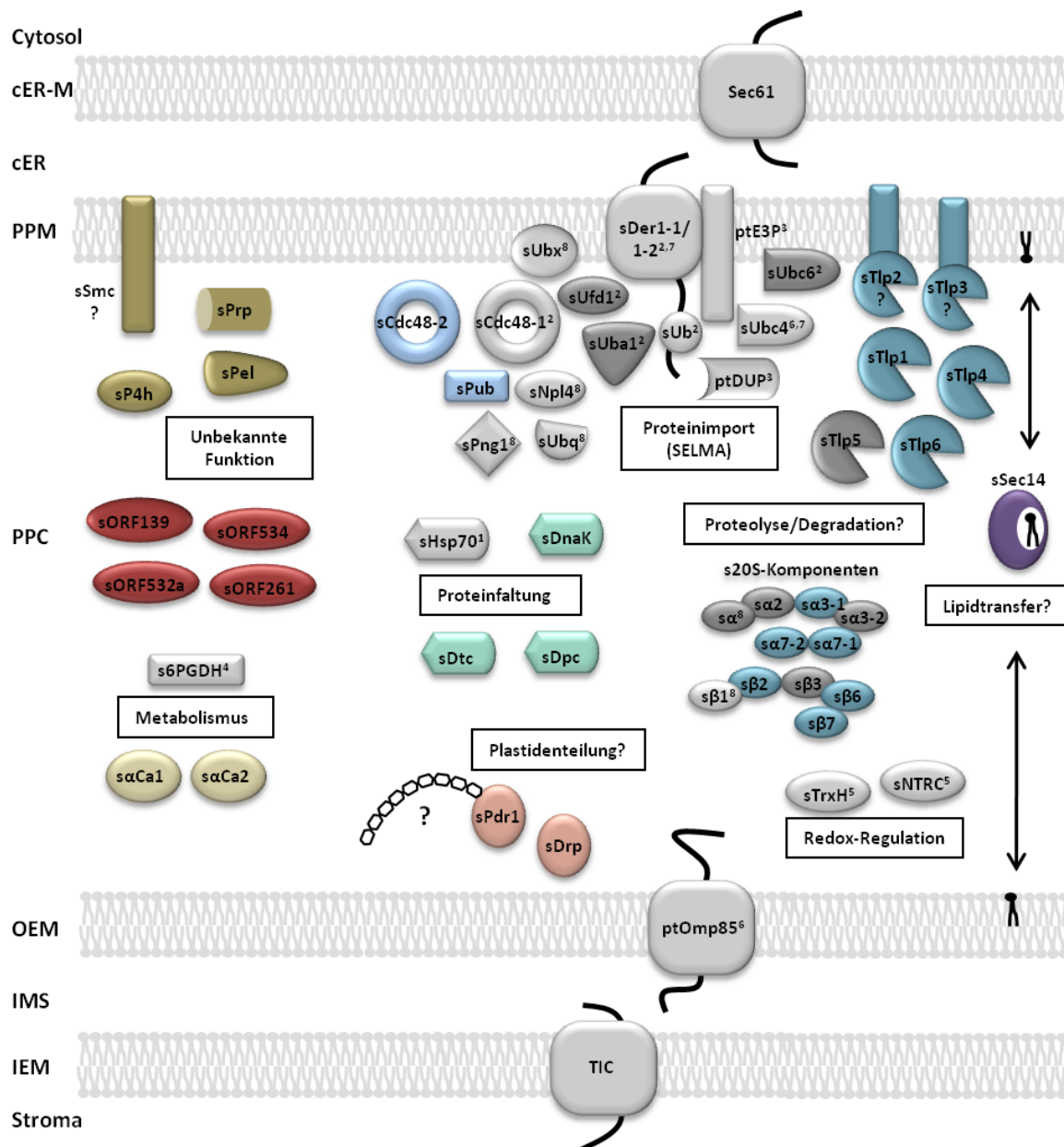


Abb. 4.2 Das PPC der Diatomee *Phaeodactylum tricornutum*. Die in dieser Arbeit identifizierten und verifizierten PPC-Proteine sind farblich hervorgehoben und nach putativen Funktionen (schwarze Rahmen) kategorisiert. Solche Proteine die bisher nur *in silico* als potentiell PPC-lokalisiert eingestuft wurden, sind in dunkelgrau dargestellt. PPC-Faktoren, die bereits in anderen Studien identifiziert wurden sind in hellgrau (verifiziert) bzw. dunkelgrau (*in silico*) abgebildet und entsprechend gekennzeichnet. ¹, Gould et al. (2006); ², Sommer et al. (2007); ³, Hempel et al. (2010); ⁴, Gruber et al. (2009); ⁵, Weber et al. (2009); ⁶, Bullmann et al. (2010); ⁷, Hempel et al. (2009); ⁸, Stork et al. (2012). cER(-M), chloroplast ER(-Membran); PPM, periplastidäre Membran; PPC, periplastidäres Kompartiment; OEM, plastid outer envelope membrane; IMS, intermembrane space; IEM, plastid inner envelope membrane.

Ein Großteil der PPC-residenten Proteine ist somit wahrscheinlich in die Homöostase der ehemals primären Plastide (Proteintransport, Lipidtransfer, Plastidenteilung), sowie des PPC selbst involviert. Hinsichtlich Letzterem kommen neben putativen Funktionen in Proteinfaltung (Chaperone), -umsetzung (Proteasen und proteasomale 20S Komponenten) und Redox-Regulation (sTrx, sNTRC), mögliche strukturelle Prozesse in Frage (sORFs, sPrp, sPel, sSmc, sP4h).

Zusammenfassend deuten die Resultate dieser Arbeit darauf hin, dass es sich beim PPC der Diatomee *P. tricornutum* um ein auf natürlichem Wege minimiertes Cytoplasma mit drastisch eingeschränkten metabolischen und zellulären Funktionen handelt, die wahrscheinlich vorwiegend auf die Homöostase der primären Plastide abzielen. Auf Basis der bisherigen Daten kann nicht abgeschätzt werden, zu welchem Teil die PPC-Proteinkomposition bis dato aufgeklärt wurde und welche weiteren Faktoren zu erwarten wären. Möglicherweise könnte die Fortsetzung des nicht gerichteten Ansatzes (3.1) weitere PPC-Funktionen offenbaren. Außerdem ist die Analyse der Proteinkomposition der PPM vor allem im Hinblick auf die Erlangung weiterer Erkenntnisse zu metabolischen Prozessen im PPC und einer generellen Kommunikation zwischen Wirt und Symbiont (siehe 4.5) von größtem Interesse. Hinsichtlich der bioinformatischen Identifizierung von PPC/PPM-Proteinen gilt es jedoch in zukünftigen Studien vor allem eine präzisere *in silico* Detektion von TPL-Sequenzen voranzutreiben (siehe 4.4).

4.3 Das PPC von *P. tricornutum* im evolutionären Kontext

Die Chromalveolatenhypothese, 1999 formuliert von Cavalier-Smith, beschreibt einen monophyletischen Ursprung für Organismen mit komplexen Plastiden rhodophytischer Herkunft (Cavalier-Smith, 1999). Die zunehmende Verfügbarkeit genomischer Daten dieser Protistenlinien und daraus hervorgehende phylogenetische Analysen werfen jedoch begründete Zweifel hinsichtlich der Entstehung der „Chromalveolaten“ über ein singuläres Ereignis sekundärer Endocytobiose auf. Während phylogenetische Hinweise auf eine Monophylie der Plastiden von „Chromalveolaten“ bestehen (Dorrell & Smith, 2011; Janouskovec et al, 2010; Le Corguille et al, 2009; Yoon et al, 2002), deuten Analysen mit nucleuskodierten Genen der Wirtszellen auf ein komplexeres evolutionäres Szenario hin, das mehrere unabhängige sekundäre Endosymbiosen (Burki et al, 2012b; Burki et al, 2007) oder

sogar tertiäre Endosymbioseereignisse mit Beteiligung divergenter Wirtslinien vorschlägt (Baurain et al, 2010; Bodyl et al, 2009; Sanchez Puerta & Delwiche, 2008).

Im Rahmen dieser Arbeit wurde die Proteinkomposition des perplastidären Kompartiments der Heterokontophyte *P. tricornutum* untersucht. Mittels Datenbankabgleichen mit Genomen weiterer Vertreter von Organismengruppen mit komplexer Plastide rhodophytischen Ursprungs wurde im Folgenden analysiert ob Gemeinsamkeiten in der putativen PPC-Proteinzusammensetzung dieser Organismen und *P. tricornutum* bestehen, die Hinweise auf evolutionäre Relationen liefern können. Da das PPC dieser Linien auf das ehemalige Cytoplasma einer Rotalge zurückgeht (Cavalier-Smith, 2002), wurde im Weiteren ermittelt ob sich Homologe zu PPC-Komponenten aus *P. tricornutum* in rezenten Rotalgenlinien auffinden lassen, was nicht in allen Fällen bestätigt werden konnte. (3.2).

Die *in silico* Analysen ergaben, dass einige Proteingruppen, wie z.B. SELMA-Faktoren und sORF-Proteine, in allen untersuchten Linien mit komplexer Plastide rhodophytischen Ursprungs konserviert sind (Tab. 3.1). Die generelle Konservierung der putativen periplastidären Protein-Translokationsmaschinerie konnte bereits in vorangehenden Studien beobachtet werden (Hempel et al, 2010; Sheiner et al, 2011; Sommer et al, 2007; Spork et al, 2009; Stork et al, 2012). Des Weiteren finden sich, mit Ausnahmen (Apicomplexa), u.a. auch signifikante Gemeinsamkeiten bei den PPC-lokalisierten Chaperonen, dem Thioredoxin-System und den Plastidenteilungsfaktoren (3.2). Allerdings existieren auch einige PPC-lokalisierte Proteine und Proteingruppen in *P. tricornutum*, für die sich keine Homologe in anderen Organismen mit komplexen Plastiden rhodophytischer Herkunft und Rotalgen detektieren ließen, oder die nur in nah verwandten Heterokontophyten identifiziert werden konnten (3.2).

Es wäre denkbar, dass sich die PPC-Proteinkomposition nicht ausschließlich von dem genetischen Repertoire des ursprünglichen Symbionten ableitet, sondern dass möglicherweise auch neue Faktoren und Funktionen adaptiert werden können. Dadurch könnte erklärt werden, warum sich nicht in jedem Fall Homologe zu PPC-Proteinen aus *P. tricornutum* in rezenten Rotalgengenomen auffinden lassen (siehe Tab. 3.2). Andererseits könnten diese Proteine jedoch in der ancestralen, in die sekundäre Endosymbiose involvierten Rotalge vorhanden gewesen und in den hier analysierten rezenten Linien verloren gegangen sein. Es ist anzunehmen, dass die Proteinzusammensetzung der PPCs von Organismen mit komplexen Plastiden rhodophytischen Ursprungs stark mit den vorherrschenden physiologischen Anforderungen, die sich aus den divergenten Habitaten bzw. Umweltbedingungen der

jeweiligen Organismen ergeben, korreliert. Apicomplexa, wie der Malariaerreger *Plasmodium falciparum*, führen z.B. eine parasitäre, zumeist intrazelluläre Lebensweise und betreiben keine Photosynthese in ihren komplexen Plastiden (Lim & McFadden, 2010), während Diatomeen photoautotroph leben und in aquatischen Ökosystemen existieren (Armbrust, 2009). Evolutionärer Druck und die Anpassung an unterschiedliche Lebensbedingungen sind Faktoren, die sich auf die Proteinkomposition des PPCs ausgewirkt haben könnten. Dies könnte, ausgehend von einer Monophylie des Endosymbionten in den verschiedenen Organismengruppen, die z.T. starken Abweichungen hinsichtlich der Detektion von Homologen zu PPC-lokalisierten Proteinen aus *P. tricornutum* erklären. Die Präsenz gewisser Proteingruppen im PPC, die auf *P. tricornutum* und evtl. nahe Verwandte beschränkt sind (20S Proteasom, sTlps, s6PGDH, sPrp, sPel, sSmc, sP4H, etc.), reflektiert dabei höchstwahrscheinlich die selektiven Anpassungen durch Kombination vorhandener (Symbiontenfaktoren) und neuer Elemente an spezifische physiologische Gegebenheiten hinsichtlich Metabolismus und Zellbiologie.

Im Bezug auf die kontrovers diskutierten Modelle zur Evolution von Organismen mit komplexen Plastiden rhodophytischen Ursprungs, können die aus dieser Arbeit hervorgehenden Daten über das PPC sowohl mit einer monophyletischen Entstehung (Cavalier-Smith, 1999), als auch mit dem Modell einer initial sekundären Endosymbiose und der Weitergabe der rhodophytischen Plastide über zusätzliche Endosymbioseereignisse höherer Ordnung (tertiär) (Sanchez Puerta & Delwiche, 2008) in Einklang gebracht werden. Dabei müsste jedoch der individuelle Verlust und möglicherweise Zugewinn bestimmter Proteine und Proteingruppen vorausgesetzt werden. Wird das konservierte SELMA-System als phylogenetischer Marker im evolutionären Kontext angesehen (Felsner et al, 2011), wäre die Entwicklung eines solchen Translokationssystems aus der ehemaligen ERAD-Maschinerie des Symbionten im Zuge mehrerer unabhängiger sekundärer Endosymbioseereignisse (Burki et al, 2007) jedoch nur schwer vorstellbar (evolutionäres Prinzip größtmöglicher Sparsamkeit). Ergänzende phylogenetische Analysen mit weiteren konservierten Proteingruppen, wie z.B. den sORFs, könnten weitere Hinweise bezüglich des Mysteriums der Evolution von Organismen mit komplexen Plastiden rhodophytischen Ursprungs liefern.

4.4 Ist eine *in silico* Charakterisierung des PPCs möglich?

Insgesamt konnten in dieser Arbeit durch Kombination zweier *in silico* Ansätze (gerichtet und nicht gerichtet, 3.1) und *in vivo* Lokalisationsstudien von 55 PPC-Kandidaten 29 Proteine in *P. tricornutum* tatsächlich als PPC-lokalisiert verifiziert werden. Dies entspricht einer Erfolgsquote von kaum mehr als 50%. Die Ergebnisse dieser Arbeit belegen somit, dass die *in silico* Vorhersage für PPC-lokalisierte Proteine in *P. tricornutum* mit den derzeit verfügbaren Algorithmen äußerst unpräzise und ohne experimentelle Verifizierung (*in vivo* Lokalisationsstudien) wenig vertrauenswürdig ist (3.1). Während die Detektion von Signalpeptiden (SignalP 3.0) weitestgehend zuverlässig ist, scheitern aktuelle Vorhersageprogramme wie ChloroP (Vorhersage plastidärer Transitpeptide) und TargetP (Kombination aus SignalP und ChloroP mit zusätzlicher Vorhersage mitochondrieller Zielsteuerungssequenzen) vor allem bei der Identifizierung der Transitpeptid-ähnlichen Sequenzen (TPL) von Organismen mit komplexen Plastiden (siehe 8.2, Abb. 8.1 und 8.2), obwohl einige Gemeinsamkeiten mit plastidären Transitpeptiden der Archaeplastida bestehen (Abb. 8.3). Durch beide Algorithmen konnte für die in dieser Arbeit lokalisierten PPC-Proteine aus *P. tricornutum* allenfalls eine Erfolgsquote von ca. 60 % für TargetP (vorhergesagte mitochondrielle Zielsteuerungssequenzen miteinbezogen) bzw. 33,3 % für ChloroP erzielt werden (Abb. 8.1 und 8.2). Somit ist die Detektion von TPL-Sequenzen potentieller PPC-Proteine bisher weitestgehend von der Identifizierung N-terminaler Extensionen zwischen SP und putativem maturem Protein abhängig. Die Schwierigkeit dabei ist jedoch ER- und sekretorische Proteine von PPC-Proteinen bioinformatisch zu unterscheiden, da Erstere nicht zwangsläufig direkt im Anschluss an das SP konserviert sein müssen und somit fälschlicherweise eine TPL angenommen werden kann. Dies war z.B. beim gerichteten Ansatz für 13 und beim nicht gerichteten Ansatz für 2 der falsch positiven Proteine der Fall (3.1). Vor allem bei Proteinen ohne signifikante Sequenzkonservierung und unbekannter Funktion ist eine *in silico* Klassifizierung (sekretorisch oder PPC) nahezu unmöglich (z.B. nicht gerichteter Ansatz) und die tatsächliche Lokalisation kann bisher ausschließlich über *in vivo* Lokalisationsstudien ergründet werden. Die Unterscheidung von PPC- und plastidären Proteinen kann dahingegen, nach bisherigen Kenntnissen, anhand der Präsenz oder Abwesenheit einer aromatischen Aminosäure oder Leucin an der ersten Position hinter der vorhergesagten SP-Schnittstelle (plastidäres bzw. PPC-*targeting*) getroffen werden (Gruber et al, 2007). Eine weitere Problematik, vor allem im Hinblick auf eine automatisierte Vorhersage von PPC-Proteinen, ergibt sich durch die Qualität der vorhergesagten Gen- bzw.

Proteinmodelle in der *P. tricornutum* Genomdatenbank. Oft sind Proteinmodelle N-terminal unvollständig aufzufinden, wodurch eine manuelle Korrektur der korrespondierenden Genmodelle, falls gegeben, durch EST-Daten notwendig ist, da die Detektion N-terminaler Zielsteuerungssequenzen nur mit authentischen N-Termini erfolgen kann. Dies kann bei einer automatisierten *in silico* Analyse dazu führen, dass ein signifikanter Anteil potentieller PPC-Kandidaten nicht berücksichtigt und als „falsch negativ“ eingestuft wird.

Da die manuelle Klassifizierung der subzellulären Lokalisation von Proteinen aus Organismen mit komplexen Plastiden jedoch höchst zeitaufwendig ist, bleibt es daher das Ziel die Detektion von PPC-Proteinen möglichst akkurat und verlustfrei durch bioinformatische Algorithmen zu automatisieren. Die Leistungsfähigkeit von Vorhersageprogrammen für Organismen-spezifisches subzelluläres *targeting* ist stark von dem angewandten Trainingsdatensatz abhängig. Dies wird insbesondere hinsichtlich der TPL-Vorhersage für Organismen mit komplexen Plastiden durch ChloroP bzw. TargetP deutlich, welche in erster Linie auf die Detektion von Transitpeptiden von Organismen mit primären Plastiden ausgelegt sind. Der in dieser und anderen Studien generierte Datensatz an PPC-Proteinen in *P. tricornutum* (Gruber et al, 2009; Hempel et al, 2010; Moog et al, 2011; Sommer et al, 2007; Stork et al, 2012; Weber et al, 2009) stellt somit eine solide Basis für die Entwicklung eines Vorhersagealgorithmus für PPC-spezifische TPL-Sequenzen bzw. periplastidäres *targeting* in *P. tricornutum* und verwandten Organismen dar.

Erste Ansätze zur Entwicklung eines solchen Algorithmus wurden bereits in Kollaboration mit der Gruppe um Prof. Dr. Eyke Hüllermeier vom Fachbereich Mathematik und Informatik der Philipps-Universität Marburg gestartet (Mernberger et al, 2011). Dabei konnte bisher, in Bezug auf die PPC-Proteine des Trainingsdatensatzes (siehe oben), eine Klassifizierungsgüte von ca. 90% durch einen auf *time series metrics* basierenden Algorithmus erzielt werden, was die Leistung anderer Vorhersagealgorithmen wie z.B TargetP hinsichtlich der Identifizierung PPC-spezifischer TPL-Sequenzen weit übertrifft (Mernberger et al., unveröffentlicht). Angewandt auf die über 2000 Proteinmodelle aus der Genomdatenbank bei denen ein SP durch SignalP identifiziert werden konnte (siehe 6.2.5) konnten ca. 300 Proteine als potentiell PPC-lokalisiert eingestuft werden. Vorläufige Analysen deuten darauf hin, dass sich mindestens 50% der in dieser Arbeit im PPC lokalisierten Proteine unter diesen finden. Generell wäre somit eine *in silico* Charakterisierung der Proteinkomposition des PPC in Zukunft vorstellbar, bislang bleibt jedoch eine Kombination von *in silico* Analysen und *in vivo* Lokalisationsstudien unverzichtbar.

4.5 Triosephosphat/Phosphat Translokatoren als metabolische Verbindung zwischen Wirt und Symbiont?

Die metabolische Verbindung zwischen Wirt und Symbiont gilt als eine Grundvoraussetzung der Entstehung primärer Plastiden im Zuge der primären Endosymbiose (Weber et al, 2006). Es wird vermutet, dass ein Nukleotidzuckertransporter (NST, *nucleotide sugar transporter*) aus dem Endomembransystem des Wirtes früh nach dem Ereignis der primären Endosymbiose in die innere Membran des Symbionten rekrutiert wurde und sich davon die heute existenten plastidären Phosphattranslokatoren (TPTs, PPTs, GPTs/XPTs) ableiten (Colleoni et al, 2010). Letztere transportieren phosphorylierte C3-, C5- oder C6-Zucker im Zuge eines strikten Antiportmechanismus gegen anorganisches Phosphat (Weber & Linka, 2011). Während bei primären Plastiden solche Translokatoren in der IEM präsent sind um eine metabolische Verknüpfung von Wirt und Symbiont zu gewährleisten, stellt sich die Situation in Organismen mit komplexen Plastiden wesentlich schwieriger dar, da Metabolite zwei zusätzliche Membranen (PPM und cER-Membran) überqueren müssen um im Cytosol der Wirtszelle metabolisch integriert werden zu können. Im Gegensatz zu rezenten Rotalgen, welche sowohl plastidäre TPTs als auch PPTs besitzen, konnten in Organismen mit komplexen Plastiden rhodophytischen Ursprungs bisher ausschließlich putativ plastidäre Translokatoren der TPT-Familie identifiziert werden, für die ein monophyletischer Ursprung postuliert wird (Weber et al, 2006).

Im Rahmen dieser Arbeit wurde die Präsenz von TPTs in der komplexen Plastide der Diatomee *P. tricornutum* untersucht. Dabei gelang es über *in silico* Analysen insgesamt 9 putative TPTs im Genom der Diatomee zu identifizieren (siehe 3.3). Lokalisationsstudien basierend auf Präsequenz- bzw. FL-eGFP Fusionsproteinen in Kombination mit dem Self-Assembling-GFP System ergaben, dass fünf dieser Proteine (TPT1-4, TPT6) höchstwahrscheinlich in der komplexen Plastide von *P. tricornutum* lokalisiert sind, wobei für die anderen vier eine Lokalisation außerhalb der komplexen Plastide beobachtet werden konnte (3.3). Während für TPT3 eine Lokalisation in der (c)ER-Membran mit Orientierung des C-Terminus zum Cytosol nachgewiesen werden konnte, resultierte sowohl die Expression der putativ PPC-spezifischen BTS als auch der FL-Variante von TPT2 fusioniert an eGFP in einer PPC-charakteristischen Fluoreszenzstruktur (3.3.1). TPT2 scheint demnach in der PPM der komplexen Plastide zu lokalisieren. Allerdings konnte die Orientierung von TPT2 in der

PPM nicht eindeutig geklärt werden, da es nach Anwendung des Self-Assembling-GFP Systems gleichsam zu einer Fluoreszenzausbildung mit dem ER- und PPC-Marker kam (3.3.2, Abb. 3.22). Dies könnte möglicherweise der hohen Affinität der GFP-Fragmente (S11 und S1-10) zueinander geschuldet sein, wodurch es vermutlich teilweise zur Ausbildung von Transportintermediaten kommt, was bereits in ähnlichem Kontext für Proteine der PPM (sDer1-1/1-2) in früheren Studien beobachtet werden konnte (Hempel et al, 2009). Es wäre weiterhin denkbar, dass die C-terminale S11-Fusion an TPT2_FL die akkurate Integration des Fusionsproteins in die Membran beeinflusst, zumal bei der Erzeugung der Konstrukte auf separierende *spacer* verzichtet wurde. Dennoch ist aufgrund der beobachteten Lokalisation der BTS- und FL-eGFP Fusionsproteine (PPC bzw. PPM) eine Lokalisation von TPT2 in der PPM höchstwahrscheinlich (Abb. 4.3).

TPT1 und TPT4 lokalisieren dahingegen in der inneren Membran (IEM) der komplexen Plastide von *P. tricornutum*, wobei der C-Terminus der Proteine offensichtlich dem Stroma zugewandt ist (Abb. 4.3). Ambivalente Resultate wurden bei den Lokalisationsstudien zu TPT6 erhalten (3.3.1). Während die Expression der Präsequenz von TPT6 fusioniert mit eGFP eine eindeutige ER-Lokalisation ergab, konnte für das FL-eGFP Fusionsprotein eine PPC-charakteristische, „Blob-ähnliche“ Fluoreszenzstruktur beobachtet werden, was auf eine Lokalisation des Fusionsproteins in der PPM hindeutet. Für TPT6 wurde *in silico* eine Präsequenz von 22 AS Länge identifiziert, von der das putative SP 17 AS lang ist und bei der die erste Position nach der vorhergesagten SP-Schnittstelle von einem Phenylalanin eingenommen wird (putatives „ASA-FAP-Motiv“; 3.3.1). Letzteres könnte im Kontext einer möglichen TPL-Sequenz auf ein plastidäres *targeting* (IEM/OEM) hindeuten (Gruber et al, 2007). Allerdings resultierte sowohl die Expression einer 38 als auch 63 AS langen Präsequenzvariante von TPT6 vor eGFP (3.3.1) in einer typischen ER-Lokalisation für die jeweiligen Fusionsproteine, weshalb sowohl plastidäres (primäre Plastide) als auch PPC/PPM-*targeting* für TPT6 eher unwahrscheinlich ist. Eine ergänzende Analyse der Lokalisation von TPT6 mit dem Self-Assembling-GFP System resultierte, ähnlich wie bei TPT2, sowohl in einer Fluoreszenzausbildung von TPT6_FL-S11 mit dem ER- als auch PPC-Marker fusioniert an GFP_S1-10 (3.3.2). Allerdings stellte sich die Fluoreszenz in Kombination mit dem ER-Marker nicht wie bei TPT2 in einer „Blob-ähnlichen“ (PPC), sondern in einer ER-spezifischen Fluoreszenzstruktur ähnlich zu TPT3 dar. Letzteres und die Lokalisation der Präsequenz-eGFP Fusionsproteine von TPT6 deuten somit auf eine Lokalisation im (c)ER bzw. der (c)ER-Membran hin (Abb. 4.3). Letztlich bedarf es jedoch

weiterer Studien um die tatsächliche Lokalisation von TPT6 zu klären, zumal auch eine duale Lokalisation ((c)ER-Membran und PPM) bisher nicht ausgeschlossen werden kann.

Während sich für TPT6 über Blast-Analysen vorwiegend Homologe in Pflanzen und Grünalgen finden, sind Homologe zu TPT1-4 in Rotalgen und Organismen mit komplexen Plastiden rhodophytischen Ursprungs hoch konserviert (3.3.1). Es kann daher postuliert werden, dass es sich bei Letzteren um vom ursprünglichen Rotalgenendosymbionten abgeleitete Proteine handelt, welche vermutlich über Genduplikationen entstanden sind und durch Akquisition relevanter Zielsteuerungssignale (3.3.1) von Seiten des Wirts in die unterschiedlichen Membranen der komplexen Plastide rekrutiert wurden. Phylogenetische Analysen stärken diese Hypothese (Price et al, 2012; Weber et al, 2006).

Wie unter 3.3.3 dargestellt, konnte für die TPTs aus *P. tricornutum* mit vermutlichem Rotalgenursprung (TPT1-4), auf Basis von Carbonatextraktionsstudien mit FL-eGFP Fusionsproteinen, nachgewiesen werden, dass es sich bei diesen tatsächlich um membranintegrale Proteine handelt. Allerdings konnte auch ein signifikanter Anteil an Fusionsproteinen in der löslichen Proteinfraction detektiert werden, was möglicherweise der starken Überexpression geschuldet ist oder aus einer teilweisen sterischen Hinderung der Membranintegration der putativen Tranlokatoren durch Fusion mit eGFP resultiert.

Anhand der in dieser Arbeit generierten Daten, lässt sich ein Modell hinsichtlich der Lokalisation putativer TPTs in den Membranen der komplexen Plastide von *P. tricornutum* ableiten (Abb. 4.3): Energiereiche Metabolite in Form von Triosephosphaten, die im Rahmen der Photosyntheseleistung im Symbionten generiert werden (Calvin-Zyklus), passieren die IEM wahrscheinlich über TPT1 bzw. TPT4. Der Transport von Triosephosphaten über die OEM der komplexen Plastide findet vermutlich in einem zu primären Plastiden homologen, weniger selektiven Prozess über β -barrel-Proteine wie OEP21 statt (Duy et al, 2007). Experimentelle Hinweise diesbezüglich liegen bisher jedoch nicht vor. Es wäre zudem möglich, dass TPT1 oder TPT4 evtl. dual sowohl in der IEM als auch in der OEM lokalisieren um den Transport von Metaboliten zu gewährleisten. Der weitere Transport aus dem PPC über die PPM ins cER-Lumen erfolgt höchstwahrscheinlich über TPT2. Daraufhin könnten Metabolite über TPT3 und möglicherweise TPT6 aus dem cER-Lumen in das Cytosol der Wirtszelle gelangen wo diese schließlich zur Energiespeicherung verwendet werden können,

oder ggf. über weitere putative TPTs (TPT5, TPT7-9) in andere Wirtskompartimente transportiert werden.

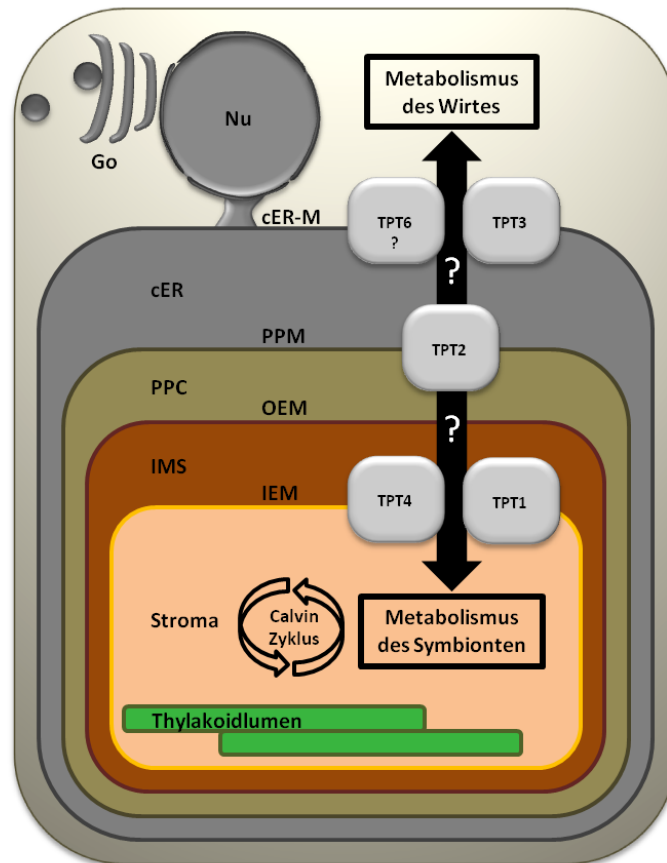


Abb. 4.3 Modell zur Lokalisation putativer Triosephosphat/Phosphat-Translokatoren in der komplexen Plastide von *P. tricornutum*. Zur Erläuterung siehe Text. cER(-M), *chloroplast ER(-Membran)*; PPM, *periplastidäre Membran*; PPC, *periplastidäres Kompartiment*; OEM, *plastid outer envelope membrane*; IMS, *intermembrane space*; IEM, *plastid inner envelope membrane*; Go, Golgi Apparat; Nu, Nucleus.

Zukünftige Studien müssen sich vor allem mit der Untersuchung der Substratspezifitäten der putativen TPTs der komplexen Plastide aus *P. tricornutum* beschäftigen, da diese nicht allein anhand der Primärsequenzen der Proteine ermittelt werden können (Lim et al, 2010; Linka et al, 2008). Es ist jedoch wahrscheinlich, dass TPT1-4 vergleichbare Substratspektren wie ihre Homologen aus Cryptophyten (Haferkamp et al, 2006) und Apicomplexen (Brooks et al, 2010; Lim et al, 2010) aufweisen (siehe 1.4) und somit die metabolische Kommunikation zwischen Wirt und Symbiont vorwiegend durch den Transport von Triosephosphaten über die Membranen der komplexen Plastide gewährleisten (Abb. 4.3).

5. Zusammenfassung

Diatomeen besitzen zwei divergente eukaryote Cytoplasmen: (i) das Cytosol der Wirtszelle, und (ii) das periplastidäre Kompartiment (PPC), welches das reduzierte Cytoplasma eines rhodophytischen Endosymbionten (komplexe Plastide) darstellt. Letzteres findet sich bei einer Vielzahl ökologisch höchst relevanter Mikroalgen und pathogener Protisten (Apicomplexa), welche komplexe Plastiden mit vier Hüllmembranen besitzen. Über die Proteinzusammensetzung und Funktion des PPC war bisher nur wenig bekannt. Im Rahmen dieser Arbeit wurde die Proteinkomposition des PPC der Diatomee *Phaeodactylum tricornutum* über einen kombinierten *in silico/in vivo* Ansatz untersucht. PPC-Kandidaten wurden dabei anhand der Präsenz spezifischer N-terminaler Zielsteuerungssequenzen genomisch identifiziert und über die Expression von eGFP-Fusionsproteinen auf ihre subzelluläre Lokalisation untersucht. Insgesamt konnten dadurch in dieser Arbeit 29 neue PPC-Proteine detektiert werden, welche sowohl bereits bekannte Funktionen im PPC, wie Proteintransport (SELMA), Proteinfaltung und Metabolismus, ergänzen, als auch auf bisher unbekannte putative Prozesse im PPC hindeuten. Insbesondere sind dabei Proteolyse und Degradation, Plastidenteilung, Lipidtransfer und eine potentielle strukturelle Organisation zu nennen. In gleichem Maße relevant sind Prozesse, die im Zuge bioinformatischer Analysen nicht im PPC der Diatomee detektiert werden konnten, was für klassische eukaryote Cytoskelett- und Vesikelkomponenten, sowie typische Signaltransduktionsfaktoren zutrifft. Die Ergebnisse dieser Arbeit legen nahe, dass es sich beim PPC von *P. tricornutum* um ein auf natürliche Weise minimiertes eukaryotes Cytoplasma mit stark begrenzten metabolischen und zellulären Kapazitäten handelt, welches mit einem autonomen eukaryoten Cytoplasma nur wenige Gemeinsamkeiten aufweist und dessen Funktion höchstwahrscheinlich vor allem der Selbsterhaltung und Homöostase der ehemals primären Plastide dient.

Ein weiteres Thema dieser Arbeit beschäftigte sich mit der Identifizierung und Lokalisation von Metabolitranslokatoren in *P. tricornutum*, die eine Verbindung zwischen den Stoffwechselprozessen von Wirt und Symbiont ermöglichen. Dabei gelang es mittels *in silico* Analysen und anschließenden *in vivo* GFP-Lokalisationsstudien mehrere putative Triosephosphat/Phosphat Translokatoren in den Membranen der komplexen Plastide von *P. tricornutum* nachzuweisen. Die Ergebnisse deuten darauf hin, dass diese, möglicherweise durch den Transport von Triosephosphaten, eine metabolische Kommunikation zwischen Wirt und Symbiont gewährleisten könnten.

5.1 Abstract

Diatoms maintain two evolutionary different cytoplasms: (i) the cytosol of the host cell and (ii) the periplastidal compartment (PPC), which resembles the reduced cytoplasm of a red algal endosymbiont (complex plastid). The PPC is present in a multitude of ecological highly relevant microalgae and pathogenic protists (Apicomplexa) that contain complex plastids surrounded by four membranes. Knowledge about the protein composition and function of the PPC as yet was only sparse. Therefore in this study the protein composition of the PPC of the diatom *Phaeodactylum tricornutum* was analyzed using a combined in silico/in vivo approach. PPC candidates were identified in the genome by means of specific N-terminal targeting sequences followed by expression of eGFP fusion proteins to investigate their subcellular localizations. This screen resulted in the identification of a total of 29 new PPC proteins, including further proteins for already known functions like protein transport (SELMA), protein folding and metabolism, as well as factors indicating new functions in the PPC. These cover protein degradation and processing, plastid division, lipid transfer and a potential structural organization. Equally important were processes genomically not detectable in the PPC, which was true for typical eukaryotic vesicular and cytoskeleton components as well as proteins for classical signal transduction. The results of this work indicate that the PPC is a naturally minimized eukaryotic cytoplasm with reduced metabolic and cellular capacities showing less similarity to an autonomous eukaryotic cytoplasm. Furthermore the results suggest that the PPC is mainly involved in maintenance of the former primary plastid and its own integrity.

A further topic of this study dealt with the identification and localization of metabolite translocators that might generate a link between metabolic symbiont and host pathways. A combination of in silico analyses and in vivo localization studies succeeded in detection of several putative triosephosphate/phosphate translocators in the membranes of the complex plastid of *P. tricornutum*. The results indicate that these proteins could facilitate a metabolic communication of host and symbiont, possibly by transport of triosephosphates.

6. Material und Methoden

6.1 Material

6.1.1 Chemikalien

Alle in dieser Arbeit verwendeten Chemikalien wurden nach Herstellerangaben eingesetzt und gelagert.

Acrylamid/Bisacrylamid		Roth
Agar Agar		Roth
Agarose NEEO		Roth
Amidoschwarz 10B		Roth
Ampicillin Trihydrat	Amp	Sigma
APS (Ammoniumpersulfat)		Roth
Beta-Mercaptoethanol		Roth
Calciumchlorid	CaCl	Roth
Cetyltrimethylammoniumbromid	CTAB	Sigma
Chloroform	CHCl ₃	Merck
Ethanol	EtOH	Roth
Ethidiumbromid	EtBr	Sigma
Harnstoff (Urea)		Roth
Isopropanol		Roth
Kanamycin	Kan	Sigma
Luminol		Roth
Phenol, TE gepuffert		Roth
Roti-Safe GelStain		Roth
Spermidin		Fluka
Trichloressigsäure	TCA	Roth
TEMED (Tetramethylethylendiamin)		Roth
Tropic marin Meersalz		Dr. Biener GmbH
Zeozin		InvivoGen

6.1.2 Enzyme

DNAseI	MBI-Fermentas
Phusion High Fidelity DNA-Polymerase F-530L (500U/μl)	Finnzymes
PNGase F (5U/μl)	Sigma
Restriktionsendonukleasen (10U/μl)	Fermentas/NEB
Revert Aid M-MuLV – Reverse Transkriptase (200U/μl)	Fermentas
SuperScript II Reverse Transkriptase (200U/μl)	Invitrogen
SuperScript III Reverse Transkriptase (200U/μl)	Invitrogen
Taq DNA-Polymerase	Biotoools
T4-DNA-Ligase (5U/μl)	Fermentas

6.1.3 Molekularbiologische Kits

DYE namic ET Terminator Cycle Sequencing Kit	GE-Healthcare
CloneJET™ PCR Cloning Kit	Fermentas
JetSorb DNA Extraction Kit	GENOMED
Phusion® High Fidelity PCR Kit	Finnzymes
Pro-Q® Emerald 300 Glycoprotein Gel and Blot Stain Kit	Molecular Probes
QIAGEN® QIAfilter Midi/Maxi Kit	QIAGEN
Rapid DNA Ligation Kit	Fermentas
μMACS™ Epitope Tag Protein Isolation Kit	Miltenyi Biotec

6.1.4 Geräte

Zentrifugen:

Centrifuge 5810R (Kühlzentrifuge)	Eppendorf
Centrifuge 5417R (Kühlzentrifuge)	Eppendorf
Centrifuge 5415D (Tischzentrifuge)	Eppendorf
Mini Spin Plus (Tischzentrifuge)	Eppendorf
Micro 22 R (Kühlzentrifuge)	Hettich
Ultrazentrifuge L7-55	Beckmann

Pipetten:

Eppendorf research 100-1000 µl	Eppendorf
Eppendorf research 20-200 µl	Eppendorf
Eppendorf research 2-20 µl	Eppendorf
Eppendorf research 0,1-2 µl	Eppendorf

PCR-Cycler:

Mastercycler-Personal	Eppendorf
Mastercycler-Gradient	Eppendorf

Agarose-Gelelektrophorese:

Elektrophoresis Power Supply EPS 200	Pharmacia Biotech
Gel-Kammern verschiedener Größen	Werkstatt PUM
Pharmacia LKB GPS 200/400 Power Supply	Pharmacia Biotech
Transilluminator	PeqLab

Filter, Membranen und Photomaterialien

Nitrocellulosemembran	Macherey-Nagel
Whatman-Filterpapier 3mm	Schleicher & Schuell
FP30/ 0,2 CA - S Sterilfilter	Schleicher & Schuell
Fuji Medical X-Ray Film 30x40 cm	Fuji
Röntgenfilmentwickler	Kodak
Röntgenfilmfixierer	Kodak

SDS-Polyacrylamidgelelektrophorese

Elektrophoresis Power Supply EPS 301/601	Amersham Biosciences
TE 77 Semy-Dry Transfer Unit	Amersham Biosciences

Sonstige Geräte:

ABI 377 Sequenzierer	Applied Biosciences
Bio-Rad Biolistic PDS-1000/He Particle Delivery System	Biorad Laboratories
Certomat R Schüttler	Braun Biotech
Exsikkator	Nalgene
Frenchpress SLM-AMINCO 4-3399	SLM-AMINCO
Klimaschrank MLR 350	Sanyo

Klimakammer
KLSM Leica TCS SP2
Nanodrop Spectrophotometer ND-1000
Spektralphotometer Ultrospec 2000
Thermocycler 60
Thermocycler comfort
TRM-V Taumelschüttler
Gel-Taumelschüttler Polymax 1040
Vortexer REAX 1DR

Weiss
Leica Microsystems
PqLab
GE-Healthcare
Bio-Med
Eppendorf
MAGV
Heidolph
Heidolph

6.1.5 Vektoren

pJet1.2
Rep (pMB1), ori, bla (amp^R), eco47IR, P_{lacUV5}, T7-Prom, MCS GeneJet

NCBI accession number: EF694056

pPha-T1

amp^R, P_{fcpA}, MCS, T_{fcpB}, sh ble (Zeozin), ori NCBI accession number: AF219942

pPha_NR

amp^R, P_{NR}, MCS, T_{NR}, sh ble (Zeozin), ori NCBI accession number: JN180663

pPha_DUAL_2xNR

NCBI accession number: JN180664

amp^R, P_{NR}, MCS1, T_{NR}, P_{NR}, MCS2, T_{NR}, sh ble (Zeozin), ori

6.1.6 Antikörper

Primärantikörper:

α -GFP (goat, WB: 1:3000) Biomol
 α -RbcL (rabbit, 1:7500) Agrisera
 α -PsbD (rabbit, WB: 1:7500) Agrisera

Sekundärantikörper:

α -goat (rabbit, HRP-gekoppelt, WB: 1:10000) Sigma
 α -rabbit (mouse, HRP-gekoppelt, WB: 1:10000) Sigma

6.1.7 Organismen

Escherichia coli TOP10 Invitrogen
F- mcrA Δ (mrr-hsdRMS-mcrBC) ϕ 80lacZ Δ M15 Δ lacX74 nupG recA1 araD139 Δ (ara-leu)7697 galE15 galK16
rpsL(Str^R) endA1 λ

Phaeodactylum tricornutum CCAP1055/1 AG Kroth, Konstanz

6.1.8 Oligonukleotide (Primer)

Die Synthese der in dieser Arbeit verwendeten Oligonukleotide (Primer) erfolgte entweder durch die Firma MWG oder Sigma. Die jeweiligen Primer wurden in PCR- oder

Sequenzierreaktionen in einer Konzentration von 5 pmol/μl eingesetzt. Die Schmelztemperaturen (T_m) der Oligonukleotide in Grad Celsius (°C) wurden von den Herstellerfirmen berechnet.

Primer für putative PPC-Proteine:

Primer-Bezeichnung	5'-	-3'	T_m °C
sCdc48-2fwEco	<u>GAATTC</u> ATGAGGTTTTCTGAGGCGGCAG		75,5
sCdc48-2BTSrvBgl	<u>AGATCT</u> TCCCGGAGCGCCTCCGCCCCGTG		87,5
sCdc48-2_2NcoI fw	<u>CCATGG</u> TCAAGAAAGGCCTCTTAC		67,6
sCdc48-2_1NcoI rv	<u>CCATGG</u> CTTGTTGTGGATTGTCCAC		73,8
sCdc48-2_3PvuI fw	<u>CGATCG</u> AGCATCGGCATCTTTTG		73,8
sCdc48-2_2PvuI rv	<u>CGATCG</u> GATACCGAACCATTCTTG		71,4
sCdc48_FLBglI rv	<u>AGATCT</u> TTTCATTGCTATCGGAATCTGAAC		68,5
sCdc48-2_fwSacI	<u>GAGCTC</u> ATGAGGTTTTCTGAGGCGGCAG		77,6
sPub_fwEcoRI	<u>GAATTC</u> ATGCATGGGCTTGAGTGGCTTG		78,8
sPub_BTSrvBamHI	<u>GGATCC</u> TGTGGGGTCGAACTCGTTCGGTA		79,3
sPub_fwNotI	<u>GCGGCC</u> GCATGCATGGGCTTGAGTGGCTTG		89,6
sPub_IntBglI fw	<u>GCCACG</u> TCGGCTCAGATACAACGGCTTG		80,1
sPub_IntBglI rv	<u>GCCGAC</u> GTGGCCTCGAATACAACCTCAG		79,3
sPub_FLrvBglI,	<u>AGATCT</u> ACCCTCCTCATCCGTAGAATCG		70,4
salpha3-1-HindIII-fw	<u>AAGCTT</u> ATGAAGACACGGTCAATGG		67,8
salpha3-1-BamHI-rev	<u>GGATCC</u> TATTGGCTTGTTTTATCG		67,7
salpha7-1_fw	<u>GAATTC</u> ATGATCGGAGACATCAT		68,9
salpha7-1_rv	<u>GGATCC</u> AATTGTTGTACCGGCTC		69,9
salpha7-2_5_EcoRI	<u>GAATTC</u> ATGGCAAGGTTTCGGACACA		73,8
salpha7-2_3_BglII	<u>AGATCT</u> TGCACTCGATCCCGAGTT		70,0
sbeta2_fw	<u>GAATTC</u> ATGCCATCGTTACTCTCC-3		66,4
sbeta2_rv	<u>GGATCC</u> TCGCAACAATCGGCATCC		76,5
sbeta6_fw	<u>GAATCC</u> ATGTTAAGAGAAGAAGT		58,4
sbeta6_rv	<u>AGATCT</u> TAAATTCTTTTCTTCTT		54,7
sbeta7_5_EcoRI	<u>GAATTC</u> ATGCATTCTGGAAGGTCGTAT		69,3
sbeta7_3_BamHI	<u>GGATCC</u> CAACGCTGCAGTGCTGG		77,5
sTlp1_fwEco	<u>GAATTC</u> ATGATTCCAAAAGTCACAGCAATG		71,7
sTlp1_BTSrvBam	<u>GGATCC</u> CGCGCTCTTCCTTGTCGTCGTGTC		81,7
sTlp2_fwMunI	<u>CAATTG</u> ATGACGAAAAGGTTGGCTGTTG		73,6
sTlp2_rvBamHI	<u>GGATCC</u> ATCCGTCGTCAGGGCTATTGCC		79,7
sTlp3_fwEcoRI	<u>GAATTC</u> ATGCTGATCTGGTTGCGCTCCAG		77,9
sTlp3_rvBamHI	<u>GGATCC</u> ACCAAACCCCGCAAAGTTGCC		82,0
sTlp3_INTfwHindIII	<u>AAGCTT</u> TGTCGAATTGTATTGGCTGG		70,1
sTlp3_INTrvHindIII	<u>AAGCTT</u> GCCTTGACGGGTTGGAATGTC		76,0
sTlp4_fwEcoRI	<u>GAATTC</u> ATGTTTCTTGCTTTGTGGCTC		70,0
sTlp4_rvBglII	<u>AGATCT</u> GTCGTGTTTATCTTTTTTCTG		63,5
sTlp4_INTfwSacI	<u>GAGCTC</u> GACGGCATCCCATACCAAGTG		77,5
sTlp4_INTrvSacI	<u>GAGCTC</u> GCGAATTGCGGGTCCACCCGAG		84,9
sTlp6_fwEcoRI	<u>GAATTC</u> ATGCGTTTGACGATGAATATACC		69,8
sTlp6_rvBglII	<u>AGATCT</u> AAGCGCATTCACCTCAAATTGC		70,2
sTlp6_INTfwXbaI	<u>TCTAGA</u> TATCGTCTTCCAGTACGATGC		66,1
sTlp6_INTrvXbaI	<u>TCTAGA</u> GTGGCCGTCTTTGATTGTGGG		73,8
sDtc_fwXbaI	<u>TCTAGA</u> ATGAGGACAATTCCTTTTCAGTTC		62,7
sDtc_rvPagI	<u>TCATGA</u> AGTTGAAGCGGTAGTGCACTCGTG		68,1
sDpc_fwEcoRI	<u>GAATTC</u> ATGTTTCGAGCGACCGATGTGCCG		69,5
sDpc_BTSrvBam	<u>GGATCC</u> ATCGGTAGCCTTTTCGCGGTTGGC		82,2
sDnaK_fwMunI	<u>CAATTG</u> ATGCTGCGCTTTGTACGCCG		78,4
sDnaK_rvBamHI	<u>GGATCC</u> GCTATTCGTGCGTTTCAAACC		75,8
Rab1b_fwEcoRI	<u>GAATTC</u> ATGCGACATACGAGTCAGTATATT		67,2
Rab1b_BTSrvBamHI	<u>GGATCC</u> GACTCCGCTATCTCCAATT		71,4
Sec24_fwXbaI	<u>TCTAGA</u> ATGACCAACTTTCCCCCTCC		64,8
Sec24_BTSrvNcoI	<u>CCATGG</u> ACGGAGGTGTAGTCCGAGCATC		71,0

SybD_fwEcoRI	<u>GAATTCATGATCGTCTACGCTATGATTTG</u>	62,4
SybD_rvNcoI	<u>CCATGGAATTTCTCCCGTATTAGTGCTG</u>	66,7
SybA_fwEcoRI	<u>GAATTCATGACCATCGTGTATGCGCTTG</u>	65,1
SybA_rvNcoI	<u>CCATGGAACCTTCTTGACGCCGACAAGAG</u>	69,5
Ank5_fwEcoRI	<u>GAATTCATGCATTGGTACCAAGTCCAAA</u>	71,3
Ank5_rvBglII	<u>AGATCTAAGCTCGGAATGATCGCTTTCT</u>	70,1
sDrp_fwSacI	<u>GAGCTCATGACCATGAGGCCGAGA</u>	74,5
sDrp_BTS_BamHI_rv	<u>GGATCCTGTATGTAGCTGATTGTAGGC</u>	64,8
sDrpINT_fwHindIII	<u>AAGCTTTCGATCGTGAATTGCAAGCGC</u>	78,9
sDrpINT_rvHindIII	<u>AAGCTTTGAGTTTTGCTTCGGTCGAGG</u>	73,6
sDrpFL_rvBamHI	<u>GGATCCTCGAGACGGCTTGTCGAACAC</u>	77,1
sPdr1(GT8-2)_fwEcoRI	<u>GAATTCATGCGCCAGCATTCAACTGCG</u>	78,8
sPdr1(GT8-2)_FLrvBamHI	<u>GGATCCTTGCTTTCTACTTTTAACAAC</u>	63,9
NDK3_fwEcoRI	<u>GAATTCATGTTTGCCCGTGCTGTCTTACG</u>	66,7
NDK3_FLrvBamHI	<u>GGATCCCTCATACACGGAGGCATCGACTG</u>	70,9
STK4_fwEcoRI	<u>GAATTCATGGCGACTGTAGCAGCGCAGGC</u>	70,9
STK4_rvBamHI	<u>GGATCCTAGCGCTGAGGCAAAGGACTCAG</u>	70,9
Pk-fw-EcoRI	<u>GAATTCATGGCGGACGC</u>	65,4
Pk-rv-Msc-int	<u>TGGCCATTATTGTATTCAGACGG</u>	66,4
Pk-fw-Msc-int	<u>TGGCCACCACCAGTCACAAGGGC</u>	78,0
Pk-rv-BamHI	<u>GGATCCTGTTGTGCAAGTATGGC</u>	68,5
CPF2_fwEco	<u>GAATTCATGAGCAGTAGCAGGTCCAAGC</u>	72,7
CPF2_BTSrvBam	<u>GGATCCCGACGAGGCGCGTGAGGTCGC</u>	86,5
GLR_fwEcoRI	<u>GAATTCATGCTTCGTACAGTCATGGGAT</u>	70,9
GLR_BTSrvBamHI	<u>GGATCCTTCCTTATTGTCGTCCGATGTG</u>	73,6
GltP_fwEcoRI	<u>GAATTCATGCGGTTCCGGTACGAAACCTG</u>	75,6
GltP_BTSrvBglII	<u>AGATCTTTCGTTGTCCGCGTTACTGTTG</u>	72,6
sSec14_fwSacI	<u>GAGCTCATGAGATTCCCTTGCGCCACG</u>	68,0
sSec14_rvBamHI	<u>GGATCCTACGATCGCCGCCAATTCCTGGCC</u>	73,6
salphaCa1_fwEcoRI	<u>GAATTCATGAGGCTGATTGCGATTTTCG</u>	76,0
salphaCa1_BTSrvBam	<u>GGATCCCCGACTCGATCACACATGTTTC</u>	78,5
salphaCa2_fwEcoRI	<u>GAATTCATGAACTTGTTGTCTCTGCC</u>	68,3
salphaCa2_BTSrvBam	<u>GGATCCACCTTTGCTGTAATCTAGACGAG</u>	70,6
alphaCa1_fwEcoRI	<u>GAATTCATGGTGGGTCTTCCATCAGTAC</u>	70,6
alphaCa1_BTSrvBam	<u>GGATCCCGTTTCAAAGCCGTCCGGCCAAC</u>	83,5
alphaCa2_fwEcoRI	<u>GAATTCATGCGCTCGTTTCTACTTTGGAG</u>	72,7
alphaCa2_BTSrvBam	<u>GGATCCCCAACCCAGGCAGTTTTTCGGTATC</u>	79,2
sORF139_fwEcoRI	<u>GAATTCATGACTGGTCTAAGATGTGTTTTTC</u>	62,9
sORF139_rvNcoI	<u>CCATGGATTTGGTAAAGGCTTTGCTTTTGGC</u>	66,8
sORF261_fwEcoRI	<u>GAATTCATGAAGTTGTCTTTCCGAGGGTG</u>	65,3
sORF261_rvBamHI	<u>GGATCCCTGCATATACAGCATATCGTTAC</u>	65,3
sORF532a_fwEcoRI	<u>GAATTCATGTTACATTGTAATTGGCAAC</u>	59,3
sORF532a_INTfwSpeI	<u>ACTAGTGGACGCCTTTGGACGGGAAG</u>	74,6
sORF532a_INTrvSpeI	<u>ACTAGTAAACCCCGAGCCGTGCGAGC</u>	76,0
sORF532a_rvBamHI	<u>GGATCCAAGTTCACATCGACCCGAAGC</u>	69,5
sORF534_FL1MEcoRIfw	<u>GAATTCATGATAAGGATGAATGTGTGAAAG</u>	69,9
sORF534_rvNcoI	<u>CCATGGAACGCTCAAGTTCAGCGACAATC</u>	68,1
ORF387_fwXbaI	<u>TCTAGAATGCTACGCAACAAGTACTCCAC</u>	65,3
ORF387_rvNcoI	<u>CCATGGATGACTTGCGCTCTGTGCGTGC</u>	71,0
sP4Hc_fw1_EcoRI	<u>GAATTCATGTGCCAACACACCCTTACGC</u>	75,2
sP4Hc_BTSrvBamHI	<u>GGATCCTTGCAATTGATCCGCAAAGTCAG</u>	76,6
sPel_fwEcoRI	<u>GAATTCATGCGCACCCAAGGCAGCAAAC</u>	79,7
sPel_BTSrvBgl	<u>AGATCTATTTTCGGCTTGGGAGTACACTCG</u>	71,7
Fru_fwEco	<u>GAATTCATGAGGCTGCGTACATCCATTGC</u>	75,9
Fru_BTSrvBam	<u>GGATCCGTAGGGCGATCCAGATGCTGAAG</u>	78,9
Tom70_BTSfwMunI	<u>CAATTGATGGCGCTCGGTACACGATC</u>	75,8
Tom70_BTSrvNcoI	<u>CCATGGTCTTGTCCAAGAGTTCAATTTTG</u>	71,9
sPrp_fwEco	<u>GAATTCATGAAGTTCTCCGCCTTCGTCCG</u>	75,5
sPrp_BTSrvBam	<u>GGATCCCGGACTATTTAGGTACCGTTCCG</u>	74,8
FolC_fwEcoRI	<u>GAATTCATGCGAACGAACATCGTGCGCTC</u>	79,7
FolC_BTSrvBamHI	<u>GGATCCGACATTCAGCGCGGCTGGATTG</u>	82,8

Hap_fwEcoRI	<u>GAATTC</u> ATGATCAACGTACGGAGACTC	68,3
Hap_BTSrvBamHI	<u>GGATCC</u> GAGAGTGCATAAATCTTCCGG	75,9
PGP_fwEco	<u>GAATTC</u> ATGTCACTGAGAATGGCGGGTAC	73,7
PGP_BTSrvBam	<u>GGATCC</u> GAGAGTCGGTTTGTTCGCGACG	81,0
PPG1_fwEcoRI	<u>GAATTC</u> ATGAAGTTTAATCTTCTCGTATTC	63,8
PPG1_BTSrvBam	<u>GGATCC</u> GTCATTCCAGTTTCCTCCTGCG	78,6
sSmc_fwEcoRI	<u>GAATTC</u> ATGAAGGGGTGTGATGCTGG	74,4
sSmc_BTSrvBamHI	<u>GGATCC</u> AGGTCCAATGCAAGACACTTCC	75,7
CoaE_fwEcoRI	<u>GAATTC</u> ATGCCCCCTTTGTATCGATTGC	73,9
CoaE_rvBamHI	<u>GGATCC</u> TTGGTTTGGATACCAATTACAGG	71,3
DCP_fwEcoRI	<u>GAATTC</u> ATGGCTTGGGAAGCGAATCATC	73,1
DCP_INTrvXbaI	<u>TCTAGA</u> TCGGCCTTTTCTTGATTGGCC	73,4
DCP_INTfwXbaI	<u>TCTAGA</u> AGTATTCGTGAAAGAATCGGG	66,9
DCP_rvBamHI	<u>GGATCC</u> CTCAATTAGCAAGCCTTTC	69,2
E1-like_fwEcoRI	<u>GAATTC</u> ATGAAGAAGGACCGTTCCGACC	74,6
E1-like_rvBamHI	<u>GGATCC</u> GTAGGCTAATCTATCATTTTGGC	70,6
M6_fwXbaI	<u>TCTAGA</u> ATGAAGTTTCCTTCGCAACTCGG	72,5
M6_rv_NcoI	<u>CCATGG</u> AGGACCATCTACCGTGAAGC	74,4
Trx_fwEcoRI	<u>GAATTC</u> ATGGATCTGCTTCTATTACCC	67,3
Trx_rvBamHI	<u>GGATCC</u> TGTTTGCGGGAGTTTGCCC	77,5
Phyco-fw-Xba	<u>TCTAGA</u> TGAAGTTGGTTCCTGCG	66,3
Phyco-rv-NcoI	<u>CCATGG</u> TGGCGATGGGAATAGC	73,0

Primer für putative Triosephosphat/Phosphat-Translokatoren (TPTs):

Primer-Bezeichnung	5'-	-3'	T _m °C
TPT1_fwEco	<u>GAATTC</u> ATGATGATGAAGCGAGCTTTAG		69,0
TPT1_BTSrvBamHI	<u>GGATCC</u> TACCGAAGACAGCGCCAAAGAG		76,3
TPT1_rvk_BamHI	<u>GGATCC</u> CTTTTTTTTGGACGCATTCATAGC		74,7
TPT2_fwEco	<u>GAATTC</u> ATGAGGTTTGCAGCATGGCTCG		77,1
TPT2_BTSrvBam	<u>GGATCC</u> CCGCGTGTCCAGTGTCTCCTTC		83,1
TPT2_FLrvBam	<u>GGATCC</u> TCGTTTCTTGCGAGTCTGCTTG		75,9
TPT3_fwEco	<u>GAATTC</u> ATGAGTTCCTCGACTGCGAAAG		72,6
TPT3_SPrvBamHI	<u>GGATCC</u> ATGGGCATGGATCATGTTGGCAGC		83,0
TPT3_FLrvBam	<u>GGATCC</u> AGAGGTTTTGGCCTTCTTCGCTG		78,1
TPT4_fwMunI	<u>CAATTG</u> ATGAAGGTCGCTACCACGCTAAC		73,4
TPT4_BTSfwEcoRI	<u>GAATTC</u> ATGAAGGTCGCTACCACGCTAAC		72,7
TPT4_BTSrvBamHI	<u>GGATCC</u> GGGCGATTTGGTTGCCCCACCGCG		89,6
TPT4_FLrvBam	<u>GGATCC</u> CTTCTTGGGCTTGACAAGATTG		74,6
TPT5_fwEco	<u>GAATTC</u> ATGGCGCAGACCAAAGCCAGTC		77,9
TPT5_FLrvBgl	<u>AGATCT</u> GGCCACCACCTTGCGTCCAGTTG		80,6
TPT6_fwEco	<u>GAATTC</u> ATGCTAGTCATTGCAATCTTGC		69,7
TPT6_BTSrvBglII	<u>AGATCT</u> GGATTGCAGAAAGTCCCGACTGG		77,8
TPT6_BTSrvS-BglII	<u>AGATCT</u> GGAGCCCCAACGTAGACAGCGACAT		76,9
TPT6_FLrvBgl	<u>AGATCT</u> CAGTGACGTAGTTACGGAATAGC		67,1
TPT7_fwEco	<u>GAATTC</u> ATGAGGCGACTTGGTAGCGAGG		75,5
TPT7_FLrvNco	<u>CCATGG</u> AATTTACAGCACTGGTAGGAAG		70,3
TPT8_fwEcoRI	<u>GAATTC</u> ATGTTTGGGCTCGATAAGCATC		71,9
TPT8_NEUrvBamHI	<u>GGATCC</u> GACTACTCGCTTGTTGTTCCG		75,0
TPT9_fwEcoRI	<u>GAATTC</u> ATGGATTCTGACAACGTATTAAC		65,9
TPT9_INTrvKpnI	<u>GGTACC</u> CGTCGCATCCAGAGTCGATGC		79,2
TPT9_INTfwKpnI	<u>GGTACC</u> GAGGAACACTCGATTGTTGAGC		74,1
TPT9_rvBglII	<u>AGATCT</u> TACATTGGAATCCTTGTGATCG		68,4

Primer für Self-Assembling GFP:

Primer-Bezeichnung	5'-	-3'	T _m °C
S1-10_fwSpeI	<u>ACTAGT</u> ATGGGTGGCACCAGTAGCAAAGGAG		73,5

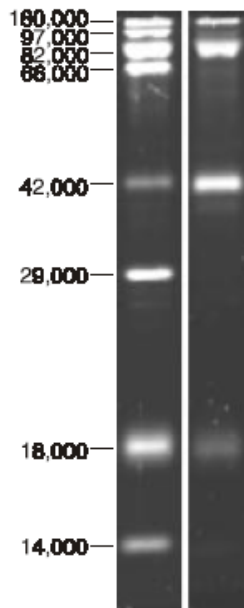
S1-10_rvSacII	<u>CCGCGGTTAGGTACCCTTTTCGTTGGG</u>	77,4
PDI_fwSpeI	<u>ACTAGTATGACTGCAATAGGGTCTTTTCGG</u>	69,0
PDI_FLrvBglII	<u>AGATCTCAATTCGCCTTCATCAAAAAGATCC</u>	72,7
sHsp70_fwSpeI	<u>ACTAGTATGGTGCATCTTCCATCCTCC</u>	68,7
sHsp70_BTSrvBamHI	<u>GGATCCTTTGTTGTCTGCATTACTGGC</u>	71,5
MGD1_fwSpeI	<u>ACTAGTATGTGCAAACTGTTTTTCC</u>	61,1
MGD1_rvBamHI	<u>GGATCCTGATGTCCGTGTCAATTTAGCTCC</u>	74,9
AtpC_fwSpeI	<u>ACTAGTATGAGATCCTTTTGCATCGC</u>	66,1
AtpC_BTSrvBamHI	<u>GGATCCTCCATCCATGACAATCGTTGC</u>	75,8

Sequenzierprimer:

Primer-Bezeichnung	5'-	-3'	T _m °C
pJet1.2_fw	GCACAAGTGTTAAAGCAGTT		57,0
pJet1.2_rv	CTCTCAAGATTTTCAGGCTGTAT		60,3
pPha-T1_MCSfw	AAGAGTGCTCGTGTGCTTCG		66,7
pPha-T1_MCSrv	GCGGCACAAATGGGCATCCTTGCTC		79,4
pPha_NR_MCSfw	GGTCGGGTTTCGGATCCTTCC		71,6
pPha_NR_MCSrv	GATGAACATAAAAACGACGATGAG		62,1
eGFP_5'rv	GGCAGCTTGCCGGTGGTGCAGATG		80,5
Dual_2xNR_MCSlrv	CTAACGCAGCTTAGACATAAAC		57,4
Dual_2xNR_MCSlrv	GTCTTATCCAGGTCCAAACAGATTG		65,4
Cdc48-1_INT3'fwSeq	GGACGTGGCGAAGCAGCGAGCAAC		79,7
48-1_Seq_fwVorHindIII	CGTTGATCTGGAAGTGTCTGGCTCG		74,6
48-1_Seq_fwWvorHindIII	CCAGACTTGAAGAGCCGTATCGC		69,9
48-1_Seq5'rv	GGTGGCAGCCCAGGACGTCGTCAGC		82,1
S1-10_Seq5'rv	CCATCTAATTCAACAAGAATTGGG		64,2
S11_rvHindIII	AAGCTTGGTACCTTATGTAATCCCAGC		67,5
S1-10_SeqKrv	GCCTTTTCATGTGATCCGGATAACGG		73,7
PDI_SeqKfw	CCTTGGAAGCGCACTAGAAATGACC		71,7
MGD_SeqKfw	GGCTTCGTCGAGAACATGGACGAATGG		77,5
MGD_SeqK2_fw	GGACGAAAATCTTCCGACCGTCTTGATCG		77,3

6.1.9 DNA- und Proteinmolekulargewichtsstandards

Die Analyse der Größe von DNA-Fragmenten auf TBE-Agarosegelen (0,8 – 2 %) erfolgte mit Hilfe der GeneRuler™ Express DNA Ladder von Fermentas bzw. der GeneDirex® 100bp DNA Ladder H3 RTU der Firma Genedirex. Die Eingesetzte Konzentration betrug 50 ng/µl. Zur Analyse von Proteinmolekulargewichten auf Polyacrylamidgelen und ggf. anschließenden Western Blots wurde die PageRuler™ Prestained Protein Ladder der Firma Fermentas verwendet. Es wurden jeweils 5µl eingesetzt. Die Analyse von Proteinmolekulargewichten und des Glycosylierungsstatus potentieller Glycoproteine auf Polyacrylamidgelen erfolgte mittels Einsatz des CandyCane™ Glycoprotein Molecular Weight Standards von Molecular Probes. Es wurden 0,5 µl Marker mit 7,5 µl PBS-Puffer (pH 7,0) und 2 µl SDS-Probenpuffer (Laemmli, 1970) versetzt, 10 min bei 95 °C erhitzt und anschließend auf ein SDS-Gel aufgetragen.



CandyCane™ Glycoprotein Molecular Weight Standards

Der Auftrennungsbereich des Glykoprotein-Molekulargewichtsstandards liegt zwischen 14 und 180 Kilodalton (kDa). Die linke Spur zeigt alle visualisierten Proteinbanden, während bei der rechten Spur nur glykosylierte Proteine mit dem Pro-Q Emerald Detektionsreagenz sichtbar gemacht wurden (Molecular Probes).

6.1.10 Hardware, Software und Internetanwendungen

Alle in dieser Arbeit erwähnten Genmodelle, sowie EST- und Sequenzierrohdaten wurden mit dem Programm Sequencer 4.7 bis 5.0 der Firma Gene Codes bearbeitet bzw. analysiert. Der Abgleich von Nukleotid- bzw. Proteinsequenzen gegen lokale Datenbanken erfolgte mittels DOS-Anwendung Blast (Altschul et al, 1997). Proteinsequenzalignments wurden mit den Programmen ClustalX (Thompson et al, 2002) und GeneDoc 2.6.002 (Nicholas & Nicholas, 1997) erstellt. Die Bearbeitung von Aufnahmen durch das Konfokale Laser Scanning Mikroskop (KLSM) wurde mit dem Programm LCS lite 2.0 der Firma Leica Microsystems durchgeführt.

Internetdienste:

Genomdatenbanken:

Phaeodactylum genome	http://genome.jgi-psf.org/Phatr2/Phatr2.home.html
Yeast genome	http://www.yeastgenome.org/
Emiliana genome	http://genome.jgi-psf.org/Emihu1/Emihu1.home.html
Thalassiosira genome	http://genome.jgi-psf.org/Thaps3/Thaps3.home.html
Aureococcus genome	http://genome.jgi-psf.org/Auran1/Auran1.home.html
Ectocarpus genome	http://bioinformatics.psb.ugent.be/webtools/bogas/overview/Ectsi
Plasmodium genome	http://plasmodb.org/plasmo/home.jsp
Toxoplasma genome	http://www.toxodb.org/toxo/home.jsp
Guillardia genome	http://genome.jgi-psf.org/Guith1/Guith1.home.html
Cyanidoschizon genome	http://merolae.biol.s.u-tokyo.ac.jp/
Galdieria genome	http://genomics.msu.edu/galdieria/
Porphyridium ESTs	http://dbdata.rutgers.edu/data/plantae/
Calliarthron partial genome	http://dbdata.rutgers.edu/data/plantae/
Cyanobase	http://genome.kazusa.or.jp/cyanobase/

Tair	http://www.arabidopsis.org/
BLAST (NCBI)	http://www.ncbi.nlm.nih.gov/BLAST

Proteintools:

NCBI-CDS	http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
SignalP3.0	http://www.cbs.dtu.dk/services/SignalP-3.0/
TargetP	http://www.cbs.dtu.dk/services/TargetP/
ChloroP	http://www.cbs.dtu.dk/services/ChloroP/
NetNGlyc1.0	http://www.cbs.dtu.dk/services/NetNGlyc/
PeptideMass	http://web.expasy.org/peptide_mass/
TMHMM	http://www.cbs.dtu.dk/services/TMHMM/
PSORT	http://psort.ims.u-tokyo.ac.jp/form.html
ProtParam	http://web.expasy.org/protparam/
SMART	http://smart.embl-heidelberg.de/smart/set_mode.cgi

Sonstige Internetdienste:

Expasy	http://expasy.org/tools/
NCBI	http://www.ncbi.nlm.nih.gov
NCBI PubMed	http://www.ncbi.nlm.nih.gov/pubmed/
WebLogo	http://weblogo.berkeley.edu/

Letzter Zugriff auf alle angegebenen URLs: 26.06.2012

6.1.11 Puffer und Lösungen

Alle in dieser Arbeit verwendeten Puffer und Lösungen sind unter der jeweiligen Methode, bei denen diese zum Einsatz kamen, aufgeführt (6.2).

6.2 Methoden

6.2.1 Kultur- und Stammhaltung

6.2.1.1 Kultur- und Stammhaltung der Diatomee *Phaeodactylum tricornutum*

Die Kultivierung von *P. tricornutum* erfolgte bei 22 °C in einem Klimaschrank bzw. einer Klimakammer unter Dauerlichteinwirkung von 8000-10000 Lux. Der Organismus wurde entweder in sterilem f/2-Flüssigmedium (pH vor dem Autoklavieren auf 7,0 mit HCl eingestellt) oder zur Stammhaltung nach Zugabe von 1,5% Agar Agar und anschließender Autoklavierung auf f/2-Festmedium in Petrischalen kultiviert. Den jeweiligen f/2-Medien wurde nach dem Autoklaviervorgang Vitamin- und Spurenelementlösung im Verhältnis 1:1000 zugesetzt.

f/2 Flüssigmedium:

Tropic marine	1,66 % (w/v)
NaNO ₃	0,89 mM, bzw.
NH ₄ Cl	1,5 mM
NaH ₂ PO ₄ x H ₂ O	36 µM
Tris (pH 8,0)	2 mM
f/2 Vitamine	1:1000
Spurenelemente	1:1000

f/2 Spurenelemente:

FeCl ₃	11,65 mM
Na ₂ EDTA	11,71 mM
CuSO ₄	39 µM
ZnSO ₄	77 µM
CoCl ₂	42 µM
MnCl ₂	910 µM
Na ₂ MoO ₄	26 µM

f/2 Vitaminlösung:

Biotin	2 µM
Cyanocobalmin	0,37 µM
Thiamin-HCl	297 µM

Abhängig von der Verwendung eines Nitrat-induzierbaren Promotors wurde den sterilen Medien Nitrat oder Ammonium als Stickstoffquelle hinzugefügt. Zur Selektion transformierter *P. tricornutum* Zellen erfolgte die Zugabe von Zeozin (Invitrogen) ins Festmedium mit einer Endkonzentration von 75 µg/ml. Die Stammhaltung positiv transformierter Klone erfolgte in Intervallen von vier bis sechs Wochen.

6.2.1.2 Kultur- und Stammhaltung von *Escherichia coli* TOP10

Zur Vervielfältigung von Plasmiden wurde der *E. coli* TOP10 Stamm der Firma Invitrogen verwendet. Die Kultivierung erfolgte auf Fest- oder in Flüssigmedium nach Zugabe eines Antibiotikums als Selektionsmarker (Ampicillin oder Kanamycin). Zur Stammhaltung positiv transformierter *E. coli* Zellen wurden 500 µl einer gut angewachsenen Kultur mit 500 µl Glycerin 1:1 gemischt und bei -80 °C gelagert.

LB-Medium:

Bacto-Trypton	1 % (w/v)
Hefeextrakt	0,5 % (w/v)
NaCl	1 % (w/v)
dH ₂ O ad 1l	
Agar Agar für LB-Platten	1,2-1,5 % (w/v)
LB-Amp 50	50 µg/ml Ampicillin
LB-Kan 10	10 µg/ml Kanamycin

6.2.2 Nukleinsäureanalytik

6.2.2.1 Isolation von Gesamt-DNA (gDNA) und -RNA (gRNA) aus *Phaeodactylum tricornutum*

Die Isolierung von Gesamt-DNA aus *P. tricornutum* erfolgte nach der CTAB-Methode (Doyle & Doyle, 1990). Gesamt-RNA wurde mit Hilfe der TRIZOL-Reagenz Methode (Invitrogen) nach Herstellerangaben isoliert. Für beide Verfahren wurde jeweils eine 150 ml Algenkultur verwendet. Die isolierte gDNA bzw. gRNA wurde in ca. 20 µl Nuclease-freiem ddH₂O aufgenommen. Die anschließende Analyse der Nukleinsäurekonzentration erfolgte mit einem NanoDrop ND-1000 Spectrophotometer der Firma Peqlab.

6.2.2.2 DNaseI-Behandlung isolierter gRNA

Um isolierte gRNA von möglichen DNA-Kontaminationen zu bereinigen wurde diese einer DNaseI-Behandlung unterzogen. Die Prozedur verlief, bei Einsatz von 1 µg gewonnener gRNA, nach Angaben der Firma Fermentas.

6.2.2.3 Reverse Transkription (RT-PCR)

Für den Einsatz isolierter und anschließend DNaseI-behandelter gRNA zur PCR, musste diese zuvor mittels Reverser Transkriptase in komplementäre DNA (cDNA) konvertiert werden (RT-PCR). Dazu wurden 50-500 ng dieser RNA entweder mit der M-MuLV-RT der Firma Fermentas oder mit der SuperScriptII/III-RT von Invitrogen nach Herstellerangaben nach Zugabe von genspezifischen bzw. Random-Hexamer-Primern umgesetzt.

6.2.2.4 Amplifikation von Nukleinsäuren - Polymerase-Kettenreaktion

Die Amplifikation spezifischer DNA-Fragmente erfolgte nach der Methode der Polymerasekettenreaktion (Mullis et al, 1986). Dazu wurde der Phusion® High Fidelity PCR Kit der Firma Finnzymes verwendet. Ein Reaktionsansatz von 25 µl beinhaltete die folgenden Komponenten: DNA-Template, 0,2 µl bzw. cDNA-Template, 1 µl; 5x Reaktionspuffer, 5 µl; MgCl₂ (50 mM), 2 µl; dNTPs (10 mM), 1 µl; Primer 1, 2 µl; Primer 2, 2 µl; Phusion-Polymerase, 0,3 µl; ddH₂O, ad 25 µl. Der Ablauf des PCR-Programms, durchgeführt in einem Mastercycler-Personal bzw. -Gradient der Firma Eppendorf, beinhaltete folgende Schritte: 1. Initiale Denaturierung, 98 °C, 30 sek.; 2. Denaturierung, 98 °C, 15 - 30 sek.; 3. Primer-Annealing, variabel (T_m -3 °C); 4. Elongation, 72 °C (variabel, ca. 15 - 30 sek./kb); 5. Finale

Elongation, 72 °C, 4 min; Abkühlung, 20 °C, 20 sek. Die Schritte zwei bis vier erfolgten, abhängig von der Art des eingesetzten DNA-Templates, in Zyklen von 29 (DNA) bis 40 (cDNA) Wiederholungen.

Kolonie-PCR Experimente zur Überprüfung der genomischen Integration von biolistisch transformierten DNA-Sequenzen in *P. tricornutum* wurden mit der Taq DNA-Polymerase von Biotools durchgeführt. Dazu wurde eine auf Zeozin-haltigem f/2-Festmedium gewachsene *P. tricornutum*-Kolonie mit einem sterilen Zahnstocher aufgenommen, in 25 µl sterilem ddH₂O gelöst und zur Freisetzung der DNA für 10 min bei 95 °C erhitzt. Der Ansatz wurde daraufhin für 5 min bei Raumtemperatur (RT) und 16100 x g zentrifugiert. Anschließend wurden 5 µl des Überstandes für die Kolonie-PCR-Reaktion eingesetzt. Ein 25 µl PCR-Ansatz beinhaltete außerdem folgende weitere Komponenten: 10x Reaktionspuffer, 2,5 µl; MgCl₂ (25 mM), 4 µl; dNTPs (10 mM), 1 µl; Primer 1, 2 µl; Primer 2, 2 µl; Taq DNA-Polymerase, 0,3 µl; ddH₂O, ad 25 µl. Das PCR-Programm wurde wie im Folgenden aufgelistet durchgeführt: 1. Initiale Denaturierung, 95 °C, 5 min; 2. Denaturierung, 95 °C, 30 sek.; 3. Primer-Annealing, variabel ($T_m - 3$ °C); 4. Elongation, 72 °C, (variabel, ca. 1 min/kb); 5. Finale Elongation, 72 °C, 5 min; Abkühlung, 20 °C, 20 sek.

Die anschließende Analyse der erzeugten Amplifikate erfolgte auf einem TBE-gepufferten 0,8 – 2 % Agarosegel (siehe 6.2.2.5).

6.2.2.5 Agarose-Gelelektrophorese und Elution von DNA Fragmenten

Die Trennung und Analyse von DNA/RNA-Fragmenten erfolgte über Gelelektrophorese auf 0,8 – 2 % (w/v) Agarosegelen in 1x TBE-Puffer bei konstanter elektrischer Spannung zwischen 150 und 210 V (Sambrook et al, 1989). Zur anschließenden Visualisierung der DNA unter UV-Licht wurde den Agarosegelen Ethidiumbromid (0,0027 % (v/v)) bzw. Roti-Safe GelStain (0,003 % (v/v)) zugesetzt.

Nach erfolgter Auftrennung wurden die jeweiligen DNA-Banden von Interesse auf einem UV-Schirm mittels Skalpell aus dem Agarosegel herausgeschnitten und die darin befindliche DNA mit dem JetSorb DNA Extraction Kit der Firma Genomed zum weiteren Gebrauch nach Herstellerangaben mit 12 bzw. 20 µl sterilem ddH₂O eluiert.

10x TBE-Puffer:		DNA-Probenpuffer:	
Borsäure	1 M	Harnstoff	4 M
Tris	0,83 M	Saccharose	50 % (w/v)
Na ₂ EDTA	10 mM	Na ₂ EDTA	50 mM
		Bromphenolblau	0,1 % (w/v)
		Xylencyanol	0,1 % (w/v)
Ethidiumbromid:	2,5 % (v/v)	RNAse A	ggf. 100 mg/ml

6.2.2.6 DNA-Restriktion

Die Restriktion von DNA-Fragmenten und Plasmid-Vektoren erfolgte mit spezifischen Restriktionsendonukleasen der Firma Fermentas bzw. NEB. Ein Standard-Restriktionsansatz enthielt eine individuell angepasste Menge an DNA, Restriktionsenzymen, sowie spezifischem Reaktionspuffer und wurde auf ein Endvolumen von 10 µl mit sterilem ddH₂O aufgefüllt. Die Restriktion erfolgte bei 37 °C über eine Inkubationszeit von mindestens 1 h. Mit dem Ansatz wurde im Folgenden, wie unter 6.2.2.5 beschrieben, verfahren.

6.2.2.7 Ligation von DNA-Fragmenten

Die Ligation von eluierten PCR-Amplifikaten in pJet1.2/blunt wurde mit Hilfe des CloneJET™ PCR Cloning Kit nach Herstellerangaben (Fermentas) durchgeführt. Ligationen von DNA-Fragmenten mit zu diesen komplementär restringierten Zielvektoren (pPha-T1, pPha_NR, pPha_DUAL_2xNR) erfolgten unter Verwendung des Rapid DNA Ligation Kit von Fermentas in einem 10 µl Ansatz nach Herstellerangaben für 30 – 60 min bei RT. Für komplexere Ligationen, bei denen mehr als drei Fragmente zusammenligiert werden sollten, wurde anstelle des Rapid 5x Ligationspuffer der 10x T4 DNA-Ligase Puffer von Fermentas in einem 20 µl Ansatz und mehrstündiger Ligation bei 16 °C verwendet. Resultierende Ligationsprodukte wurden im Anschluss in RbCl₂-kompetente *E. coli* TOP10 Zellen von Invitrogen transformiert (6.2.2.7).

6.2.2.8 Transformation von RbCl₂-kompetenten *Escherichia coli* TOP10-Zellen

Um Plasmide zu vervielfältigen, wurden diese in chemisch (RbCl₂) kompetente Zellen eingebracht (Herstellung siehe unten). Dazu wurden 50 µl RbCl₂-kompetente *E. coli* TOP10-Zellen (Invitrogen) langsam auf Eis aufgetaut (Lagerung bei -80 °C) und mit Plasmid-DNA bzw. einem Ligationsansatz gemischt. Es folgte eine Inkubation für 15- 30 min auf Eis und ein anschließender Hitzeschock für 1 min bei 42 °C im Wasserbad, wonach der Ansatz

wiederrum für 5 min auf Eis inkubiert und schließlich auf LB-Agarplatten mit geeignetem Selektionsmarker ausplattiert wurde. Die Platten wurden im Anschluss über Nacht (ü.N.) bei 37 °C inkubiert.

Herstellung Rubidiumchlorid (RbCl₂)-kompetenter *E. coli* TOP10-Zellen:

Für die Herstellung RbCl₂-kompetenter *E. coli* Bakterien wurden TOP10-Zellen von Invitrogen verwendet. Eine *E. coli* TOP10-Einzelkolonie wurde in 2 ml LB-Medium über Nacht schüttelnd bei 37 °C inkubiert. Am nächsten Tag wurden 100 ml LB-Medium mit der Übernachtskultur angeimpft und bis zu einer OD₆₀₀ von 0,3 schüttelnd bei 37 °C inkubiert. Nach Zugabe von MgSO₄ und MgCl₂ bis zu einer Endkonzentration von jeweils 10 mM wurden die Zellen erneut schütteln bei 37 °C bis zu einer OD₆₀₀ von 0,5 – 0,6 inkubiert, im Anschluss 15 – 30 min auf Eis gelagert und bei 4 °C in 50 ml Falcons pelletiert. Der Überstand wurde verworfen, das Pellet in 33 ml RF1-Puffer resuspendiert und erneut 15 - 30 min auf Eis inkubiert. Anschließend wurden die Zellen pelletiert und in 5 ml RF2-Puffer resuspendiert. Die Zellen wurden in 15 ml Gefäße überführt und 15 – 30 min auf Eis gelagert, bevor sie bei 4 °C in Aliquots zu 100 µl in 1,5 ml Reaktionsgefäße überführt, in flüssigem Stickstoff schockgefroren und bei -80 °C gelagert wurden.

RF1-Puffer:

RbCl ₂	100 mM
Mn x 4H ₂ O	50 mM
KAc	30 mM
CaCl ₂ x 2H ₂ O	10 mM
Glycerin	15 % (w/v)

RF2-Puffer:

MOPS	10 mM
RbCl ₂	10 mM
CaCl ₂ x 2H ₂ O	75 mM
Glycerin	15 % (w/v)

Der pH von RF1 wurde mit 0,2 M Essigsäure, der pH von RF2 mit NaOH auf jeweils 5,8 eingestellt bevor beide Puffer steril filtriert wurden.

6.2.2.9 Plasmidisolierung aus *Escherichia coli* mittels alkalischer Lyse

Die Isolierung von Plasmiden aus transformierten *E. coli* TOP10-Zellen erfolgte über die Methode der alkalischen Lyse (Birnboim & Doly, 1979). Hierbei wurden zunächst 1,5 ml einer Übernachtskultur pelletiert, der Überstand verworfen und das Pellet in 200 µl P1 resuspendiert. Im Anschluss wurden 200 µl P2 zugesetzt, das Reaktionsgefäß wurde mehrfach invertiert und bei RT für 5 min inkubiert. Nach Zugabe von 200 µl P3 und 50 µl Chloroform erfolgte eine weitere fünfminütige Inkubation auf Eis, bevor die Reaktionsgefäße für 10 min

bei 20000 x g und 4 °C zentrifugiert wurden. Der Überstand (wässrige Phase) wurde in ein neues Reaktionsgefäß überführt und zur Fällung der Plasmid-DNA mit einem 2/3-Volumen Isopropanol gemischt. Es folgte ein weiterer Zentrifugationsschritt für 20 min bei 20000 x g und 4 °C. Der Überstand wurde entfernt, das resultierende DNA-Pellet wurde in 400 µl 70 % Ethanol resuspendiert und erneut für 5 min bei 20000 x g und 4 °C zentrifugiert. Nach Abnahme des Überstandes wurde das Plasmid-haltige Pellet im Exsikkator getrocknet und schließlich in 50 µl ddH₂O aufgenommen.

P1-Puffer:

Tris/ HCl (pH 8,0)	50 mM
EDTA	10 mM
RNase A	100 µg/ml

P2-Puffer:

NaOH	200 mM
SDS	1 % (w/v)

P3-Puffer:

KOAc (pH 5,5)	3 M
(pH mit Essigsäure eingestellt)	

Für die biolistische Transformation von *P. tricornutum* bzw. um größere und reinere Mengen an Plasmiden zu erhalten, wurden diese mit dem QIAGEN® QIAfilter Midi/Maxi Kit aus einer transformierten 50 ml *E. coli* TOP10-Übernachtskultur nach Herstellerangaben (Qiagen) isoliert.

6.2.2.10 DNA-Sequenzierung (ABI-377)

Die Sequenzierung von generierten Plasmiden erfolgte nach der Kettenabbruchmethode nach Sanger (Sanger et al, 1977) und wurde mit dem DYE namic ET Terminator Cycle Sequencing Kit der Firma GE-Healthcare nach Herstellerangaben durchgeführt. Ein Standardreaktionsansatz von 10 µl setzte sich aus den folgenden Komponenten zusammen: DNA-Template, variabel (1 – 6 µg); Primer, 2 µl (5 pM); Sequenzierreagenz, 2 µl; ddH₂O, ad 10 µl. Die Sequenzierreaktion erfolgte nach folgendem Programm: 1. Initiale Denaturierung, 95 °C, 3 min; 2. Denaturierung, 95 °C, 30 sek.; 3. Primer-Annealing, 55 °C, 30 sek.; 4. Elongation, 60 °C, 1,5 min; Wiederholung der Zyklen 2 – 4, 29x; 5. Finale Elongation, 60 °C, 4 min; Abkühlung, 20 °C, 20 sek. Im Anschluss an die Sequenzierreaktion wurde der Ansatz mit 26 µl ddH₂O und 64 µl Ethanol (abs.) versetzt, für 20 min im dunkeln bei RT inkubiert und bei 4 °C und 20000 x g für 20 min zentrifugiert. Der Überstand wurde verworfen, es wurden 500 µl unvergällten 70 % Ethanol zugegeben und erneut für 5 min unter gleichen Bedingungen zentrifugiert. Nach Abnahme des Überstandes wurde das DNA-Pellet im

Vakuum getrocknet, in 3 µl Formamid Ladepuffer aufgenommen für 1 min wie unter vorherigen Bedingungen zentrifugiert. Die anschließende Sequenzanalyse erfolgte mit dem ABI-377 Sequenzierautomaten der Firma Applied Biosystems nach Herstellerangaben. Die Auswertung der Sequenzdaten wurde mit dem Programm Sequencher (Version 4.7 – 5.0) von Genecodes durchgeführt.

6.2.2.11 Biolistische Transformation der Diatomee *Phaeodactylum tricornutum*

Die Diatomee *P. tricornutum* wurde zur homologen Expression von Reporter-Fusionskonstrukten nach den abgewandelten Protokollen von Apt et al. (Apt et al, 1996) und Zaslavskaja et al. (Zaslavskaja et al, 2000) biolistisch transformiert. Dies erfolgte unter Verwendung von aus PCR amplifizierten und mit GFP-Reportern fusionierten Gensequenzen in pPha-T1, pPha_NR bzw. pPha_2xNR_Dual Plasmidvektoren.

Für die Vorbereitung der Microcarrier wurden 60 mg M10 Wolfram-Partikel (0,7 µm) von Biorad in einem sterilen 1,5 ml Reaktionsgefäß abgewogen. Die Microcarrier wurden mit 1 ml Ethanol (abs., HPLC-Grade) versetzt, 3 bis 5 min gemischt (Vortexer) und anschließend 1 min bei 16100 x g und RT abzentrifugiert. Nach Verwerfen des Überstandes wurde das Verfahren weitere zweimal mit sterilem ddH₂O wiederholt, bevor die Microcarrier schließlich in 1 ml ddH₂O resuspendiert, in 50 µl Aliquots aufgeteilt und bei -20 °C gelagert wurden.

Zur Transformation von *P. tricornutum* wurde eine ca. 7 Tage alte Flüssigkultur eingesetzt. Zunächst wurde die Zellzahl der Kultur mittels Thoma-Zählkammer bestimmt (Gesamtzellzahl = Ø Zellen pro Kleinquadrat x 4 x 10⁶ x Volumen der Flüssigkultur in ml). Die Kultur wurde anschließend bei 1500 x g für 5 min geerntet und mit f/2-Flüssigmedium auf eine Zellzahl von 10⁸ Zellen/100 µl eingestellt. Im Anschluss wurden jeweils 100 µl davon auf Agarplatten mit f/2-Festmedium ohne Selektionsmarker mittels Impföse kreisrund (Durchmesser ca. 5 cm) ausgestrichen und für eine Stunde antrocknen lassen.

Für die Präzipitation der Plasmid-DNA an die Wolfram-Partikel wurden die folgenden Komponenten in einem Reaktionsgefäß vereinigt: 50 µl resuspendierte M10 Wolfram-Partikel, 5 µg Plasmid-DNA in 5-10 µl ddH₂O, 50 µl 2,5 M CaCl₂ und 20 µl 0,1 M Spermidin. Der Ansatz wurde 1 min gemischt (Vortexer) und für 10 min zur Sedimentation ruhen gelassen. Der Überstand wurde entfernt und das Pellet in 200 µl Ethanol (abs., HPLC-grade) resuspendiert, erneut wie zuvor gemischt und wiederum für 5 min zur Sedimentation

inkubiert. Nach Verwerfen des Überstandes und Zugabe von 50 µl Ethanol (abs., HPLC-Grade) wurde der Ansatz direkt zur Transformation eingesetzt.

Für die biolistische Transformation wurde zunächst der Innenraum, sowie alle Bestandteile des Bio-Rad Biolistic® PDS-1000/He Particle Delivery System mit HPLC-Grade Ethanol (abs.) gereinigt (Macrocarrier, Rupture Disk, Metallring) bzw. abgeflammt (Metallnetz). 15 µl der Plasmid-DNA besetzten Microcarrier Suspension wurden auf die Mitte eines Macrocarriers pipetiert und antrocknen lassen (jeweils dreimal pro zu transformierendem Konstrukt). Der Zusammenbau der Apparatur erfolgte indem zuerst die Rupture Disk (1350 bzw. 1550 psi) im oberen Teil eingesetzt und befestigt wurde. Im Weiteren wurde die Microcarrier-Abschussvorrichtung (Microcarrier Launch Assembly), bestehend aus Halterung, Macrocarrier mit DNA-präzipitierten Microcarriern und Stopping Screen, zusammengesetzt und eingebaut. Schließlich wurde die geöffnete Zielplatte mit ausplattierter Algenkultur eingesetzt, die Kammer verschlossen und ein Vakuum von -23 bis -25 psi aufgebaut, bevor die Zellen mit den Microcarriern bei einem Druck von 1350 bzw. 1550 psi Helium bombardiert wurden. Im Anschluss an die Transformation wurden die Platten mit Parafilm versiegelt und für 24 Stunden bei Dauerlichteinwirkung (8000 – 10000 Lux) und 22 °C im Klimaschrank inkubiert. Die Zellen wurden danach mit 1 ml f/2-Medium von der Platte abgeschwemmt, zu gleichen Teilen auf drei frischen f/2-Platten mit jeweils 75 µg/ml Zeozin als Selektionsmarker ausplattiert und unter den oben genannten Bedingungen inkubiert. Nach ca. zwei Wochen wurden gewachsene Kolonien für die Konfokalmikroskopie (6.2.4) vorbereitet (6.2.3.10), bzw. zur Stammhaltung weiter überimpft (6.2.1.1).

6.2.3 Proteinanalytik

6.2.3.1 Herstellung von Proteingestamtextrakt aus *Phaeodactylum tricornutum*

Die Herstellung eines Proteingestamtextrakts aus *P. tricornutum* erfolgte ausgehend von einer ca. 7 Tage alten 50 ml Flüssigkultur. Nach Pelletieren der Kultur bei 1500 x g und 22 °C für 5 min wurde der Überstand verworfen und das Pellet in 200 µl alkalischem Lysepuffer (inklusive PIC, 1:200) resuspendiert. Es folgte eine 10 min Inkubation auf Eis bei gelegentlichem Mischen (Vortexer), bevor die Suspension mit ddH₂O auf 1 ml aufgefüllt wurde. Im Folgenden wurden die Proteine mit Trichloressigsäure (Endkonzentration 10 %) gefällt. Dazu erfolgte nach Zugabe und Mischen eine 30 min Inkubation auf Eis und ein Zentrifugationsschritt bei 20000 x g und 4 °C für 15 min, wonach mehrfach mit eiskaltem 80 % Aceton gewaschen wurde. Der ausgefällte Proteingestamtextrakt wurde im Anschluss

unter Vakuum getrocknet, in 100 µl Harnstoffpuffer aufgenommen und bei 60 °C schüttelnd für 20 min gelöst, zentrifugiert (20000 x g, 4 °C für 5 min) und der Überstand wurde bei -20 °C zur weiteren Verwendung gelagert.

Harnstoffpuffer:

Urea	8 M
Tris/HCl	200 mM
EDTA	0.1 mM
SDS	5 % (w/v)
Bromphenolblau	0,03 % (w/v)
β-Mercaptoethanol	1 % (v/v)

Lysepuffer:

NaOH	1,85 M
β-Mercaptoethanol	7,5 % (v/v)
PIC	1:200

Der Protease Inhibitor Cocktail (PIC) setzte sich aus den folgenden Komponenten mit einer Endkonzentration von jeweils 200 µg/ml in HEPES/KOH (1 M, pH 7,0) zusammen: Aprotinin, Chymostatin, Antipain, Elastatinal, Leupeptin, Pepstatin A, Trypsin-Inhibitor, Na₂-EDTA. Die jeweiligen Proteinase-Inhibitoren wurden als Stocklösungen nach Herstellerangaben angesetzt. Der fertige Protease Inhibitor Cocktail wurde vor Gebrauch sterilfiltriert.

6.2.3.2 Protein-Konzentrationsbestimmung mittels Amidoschwarz 10B

Die Konzentrationsbestimmung von Proteinextrakten bzw. -lösungen mittels Amidoschwarz 10B erfolgte nach einem abgewandelten Protokoll von Schaffner und Weissmann (Schaffner & Weissmann, 1973). Dazu wurden 2 - 5 µl einer Proteinlösung mit ddH₂O auf 100 µl aufgefüllt und mit 400 µl Amidoschwarz-Färbelösung gemischt. Der Ansatz wurde bei 20000 x g und 4 °C für 15 min zentrifugiert, der Überstand verworfen und das Pellet in 1 ml Amidoschwarz-Waschlösung resuspendiert. Es folgte eine weiterer Zentrifugationsschritt (siehe oben), wonach das Pellet unter Vakuum getrocknet, in 1 ml 0,2 M NaOH resuspendiert und die Absorption bei 615 nm mittels Photometer in Einmalküvetten bestimmt wurde. Anhand einer zuvor mit bekannten BSA-Konzentrationen erstellten Eichgerade, wurde anschließend die jeweilige Proteinkonzentration aus der Steigung der Eichgeradengleichung berechnet.

Amidoschwarz-Färbelösung:

Methanol	90 % (v/v)
Essigsäure	10 % (v/v)
Amidoschwarz 10B	0,01 % (w/v)

Amidoschwarz-Waschlösung:

Methanol	90 % (v/v)
Essigsäure	10 % (v/v)

6.2.3.3 SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE)

Die Auftrennung von Proteinen erfolgte nach der Methode der SDS-Polyacrylamid-Gelelektrophorese (Laemmli, 1970). Die Acrylamidkonzentration der in dieser Arbeit eingesetzten SDS-Gele betrug 12,5 % (Auftrennungsbereich: 25 – 100 kDa) und setzte sich zusammen aus dem Trenngel mit 4,1 ml Acrylamid/N,N'-Methylenbisacrylamid (30:0,8), 3,2 ml ddH₂O, 2,5 ml 4x Trenngelpuffer, sowie 20 µl TEMED und 150 µl APS (10 % (w/v)). Das darauf gegossene Sammelgel bestand aus 0,9 ml Acrylamid/N,N'-Methylenbisacrylamid (30:0,8), 2,8 ml ddH₂O, 1,25 ml 4x Sammelgelpuffer, sowie 15 µl TEMED und 85 µl APS (10 % (w/v)). Vor Beladen des Gels wurden die Proteinproben für 10 min bei 95 °C zur Denaturierung erhitzt. Der Gellauf erfolgte bei 120 - 150 V und 25 - 35 mA. Als Marker wurde die PageRuler™ Prestained Protein Ladder der Firma Fermentas eingesetzt.

4x Trenngelpuffer:

Tris/HCl (pH 8,8)	1,5 mM
SDS	0,4 % (w/v)

4x Sammelgelpuffer:

Tris/HCl (pH 6,8)	500 mM
SDS	0,4 % (w/v)

10x SDS-Laufpuffer:

Tris	250 mM
Glycin	2 M
SDS	1 % (w/v)

6.2.3.4 Coomassie-Färbung von SDS-Gelen

Zur Visualisierung von Proteinbanden in SDS-Gelen wurde eine Färbung mit Coomassie Brilliant Blue G-250 durchgeführt. Dazu wurde das SDS-Gel dreimal für 2 - 5 min in ddH₂O gewaschen und dabei vor jedem Waschschrift kurz in der Mikrowelle erhitzt (ohne zu sieden). Anschließend wurde die Coomassie-Färbelösung zugegeben, das Gel wurde kurz erhitzt und ü.N. auf einem Taumelschüttler inkubiert. Am nächsten Tag wurde das Gel durch mehrmaliges Waschen (und Erhitzen) mit ddH₂O entfärbt und mittels Scanner oder Digitalkamera dokumentiert.

Coomassie-Färbelösung:

Coomassie Brilliant Blue G-250	0,008 % (w/v)
HCl	35 mM

6.2.3.5 Western Blot und Immunodetektion von Proteinen

Die Methode des Western Blots nach Towbin (Towbin et al, 1979) ermöglicht es gelelektrophoretisch separierte Proteinbanden auf eine Nitrocellulose-Membran zu transferieren und anschließend mittels spezifischer Immunodetektion nachzuweisen. Dafür

wurde ein SDS-Gel nach erfolgter Separation der Proteine durch PAGE zunächst 10 min in Transferpuffer inkubiert. Als Blotting-Apparatur wurde das *TE 77 Semi-dry Transfer-System* von GE-Healthcare verwendet. Whatman-Papiere, sowie die Nitrocellulose-Membran (Macherey-Nagel) wurden vor Aufbau des Blots in Transferpuffer äquilibriert. Der Blot wurde wie folgt zusammengesetzt: Whatman-Papier, 3x; Nitrocellulose-Membran; SDS-Gel; Whatman-Papier, 3x. Der Blot wurde luftblasenfrei zusammengebaut und der Transfer erfolgte über 1 h und 10 min bei einer Stromstärke von 1 mA/cm^2 und einer Spannung von 50 V.

Für die Immunodetektion der Proteine wurde die Nitrocellulose-Membran zunächst für 1 h bei RT mit ca. 15 ml Blockingreagenz (BR) auf einem Taumelschüttler geblockt. Anschließend wurde die Membran ü.N. bei 4°C schüttelnd in 4 ml BR mit spezifischem Erstantikörper (siehe 6.1.6) behandelt. Am darauffolgenden Tag wurde die Nitrocellulose-Membran dreimal für 15 min mit 10 ml TBS-T gewaschen und anschließend für 1 h mit HRP-gekoppeltem Zweitantikörper in 4 ml BR inkubiert (6.1.6). Ferner wurde erneut dreimal mit TBS-T und jeweils einmal mit TBS und dH_2O gewaschen, bevor die spezifische Detektion des Sekundärantikörpers mittels ECL (Enhanced Chemilumineszenz) erfolgte. Letzteres wurde durch Zugabe eines geeigneten Volumens an ECL (abhängig von der Fläche der Membran), fünfminütiger Inkubation und anschließender Detektion mittels Röntgenfilm in der Dunkelkammer durchgeführt.

Transferpuffer:

Tris	25 mM
Glycin	192 mM
Methanol	20 % (v/v)
SDS	0,1 % (w/v)

Blockingreagenz:

Magermilchpulver in TBS-T	5 % (w/v)
---------------------------	-----------

TBS-T:

Tris/HCl (pH 7,5)	100 mM
NaCl	150 mM
Tween20	0,1 % (v/v)

TBS:

Tris/HCl (pH 7,5)	100 mM
NaCl	150 mM

ECL-Lösung:

Luminol in DMSO	5 mM
Coomaric Acid in DMSO	0,8 mM
Tris/HCl (pH 8,5)	200 mM
H_2O_2 (30 %)	1:1000 (kurz vor Gebrauch)

6.2.3.6 Immunopräzipitation von rekombinanten Proteinen aus

Phaeodactylum tricornutum

Die Immunopräzipitation (IP) überexprimierter rekombinanter GFP-Fusionsproteine aus *P. tricornutum* erfolgte ausgehend von einer etwa zwei Wochen alten 300 ml Flüssigkultur. Während die Expression von auf pPha-T1 kodierten Fusionsproteinen dauerhaft stattfand, wurden auf pPha_NR kodierte Proteine durch einen Medienwechsel (NO_3^- anstelle von NH_4^+ als Stickstoffquelle) lediglich für ein bis zwei Tage induziert exprimiert. Die Kultur wurde bei 1500 x g und 22 °C für 5 min geerntet und in 4 ml IP-Puffer resuspendiert. Es folgten 3 - 5 Passagen in einer French Press (Aminco) zum Zellaufschluss bei 1000 psi und eine anschließende Zentrifugation bei 10000 x g und 4 °C für 10 min um nicht aufgeschlossene Zellen und Zelldebris zu sedimentieren. Der Überstand wurde im Weiteren zur Immunopräzipitation mittels $\mu\text{MACS}^{\text{TM}}$ Epitope Tag Protein Isolation Kit der Firma Miltenyi Biotec eingesetzt. Die IP über $\alpha\text{-GFP}$ -gekoppelte magnetische *beads* wurde nach Herstellerangaben durchgeführt.

IP-Puffer:

Tris/HCl (pH 7,5)	50 mM
KOAc	200 mM
EDTA	1 mM
Glycerin	10 % (v/v)
NP-40	1 % (v/v) (kurz vor Gebrauch zugegeben)
PIC	1:200

6.2.3.7 Carbonatextraktion

Um die Membranintegration überexprimierter Triosephosphattransporter-GFP Fusionsproteine zu analysieren wurde die Methode der Carbonatextraktion angewandt. 150 ml einer ca. eine Woche alten *P. tricornutum*-Kultur wurden bei 1500 x g und 22 °C für 5 min geerntet und anschließend in 4 ml Solubilisierungspuffer A (SolA) resuspendiert. Es folgten fünf Passagen in einer French Press (Aminco) zum Zellaufschluss bei 1000 psi und die anschließende Abzentrifugation noch intakter Zellen bei 1500 x g und 4 °C für 5 min. Der Überstand (Proteingestamtextrakt) wurde im Weiteren zu jeweils 2,5 ml in geeignete Ultrazentrifugen-Röhrchen geladen und in einem TFT80.4 Rotor in einer Ultrazentrifuge (Beckmann L7-55) bei 120000 x g und 4 °C für 1 h zentrifugiert. Der Überstand (lösliche Proteine) wurde abgenommen und das Pellet wurde in 2,5 ml Carbonatpuffer resuspendiert, 30 min auf Eis inkubiert und erneut wie zuvor zentrifugiert um assoziierte Proteine und Membranproteine zu fraktionieren. Nach der Zentrifugation wurde der Überstand (assoziierte Proteine) abgenommen und das Pellet (Membranproteine) in 1,5 ml SolA resuspendiert. Im

Weiteren wurde mit den Proben, wie unter 6.2.3.1 (ab Fällung der Proteine mit TCA) beschrieben, verfahren. Nach abschließender Konzentrationsbestimmung der Proben wurden diese zu gleichen Teilen auf ein SDS-Gel geladen und mittels SDS-PAGE, Western Blot und Immunodetektion analysiert.

Solubilisierungspuffer A:

NaCl	50 mM
Imidazol/HCl (pH 7,0)	50 mM
6-Aminocapronsäure	2 mM
EDTA	1 mM
Sucrose	8,5 % (w/v)
PIC	1:200

Carbonatpuffer

NaHCO ₃ (pH 11,5)	100 mM
PIC	1:200

6.2.3.8 Färbung zum Nachweis von Glycoproteinen (*Glyco-Stain*)

Die Analyse des Glycosylierungsstatus potentieller Glycoproteine, insbesondere von sPdr1-eGFP, erfolgte mit dem Pro-Q® Emerald 300 Glycoprotein Gel and Blot Stain Kit von Molecular Probes nach Herstellerangaben. Als Marker, sowie Positiv- und Negativkontrolle, diente der CandyCane™ Glycoprotein Molecular Weight Standard von Molecular Probes (6.1.9).

6.2.3.9 PNGase F-Behandlung von Glycoproteinen

Um N-glycosylierte Proteine spezifisch zu deglycosylieren wurde das Enzym PNGase F aus *Elizabethkingia miricola* der Firma Sigma-Aldrich benutzt. Bei PNGase F handelt es sich um eine Endoglycosidase, welche spezifisch N-glycosidische Bindungen, außer N-Glycanen mit N-Acetylglucosamin- α (1-3)-Fucose Bindung, an Proteinen spaltet (Tarentino & Plummer, 1994). Die Methode erfolgte unter Verwendung des Eluats aus 6.2.3.6 (25 oder 50 μ l) und 50 mM Natriumphosphatpuffer (pH 7,5) in Lösung nach Herstellerangaben.

6.2.3.10 *In vivo* Lokalisationsstudien - Expression rekombinanter GFP-Reporterfusionsproteine in *Phaeodactylum tricornutum*

6.2.3.10.1 Präsequenz- und „volle Länge“-GFP Lokalisationsstudien

Die Analyse der Lokalisation von BTS- bzw. Präsequenzen, sowie von Proteinen in voller Länge fusioniert mit eGFP (*enhanced green fluorescent protein*) in *P. tricornutum* erfolgte mittels Konfokaler Laser Scanning Mikroskopie (6.2.4). Die genkodierenden Sequenzen wurden dazu, abhängig von der EST-Datenlage (siehe Tab. 8.1), aus genomischer DNA oder

cDNA amplifiziert und zumeist über *EcoRI/BamHI* 5' an *egfp* (*BamHI/HindIII*) fusioniert und in pPha-T1 bzw. pPha_NR (*EcoRI/HindIII*) eingebracht. In der Regel kam dabei eine modifizierte Variante von *egfp* mit 6His (Hexa-Histidin)-Tag kodierendem Bereich am 3'-Ende zum Einsatz. Die zur Amplifikation der Gensequenzen eingesetzten Primer sind unter 6.1.8 aufgeführt.

Die Expression der mit eGFP fusionierten Sequenzen *in vivo* erfolgte nach der biolistischen Transformation entweder dauerhaft (Licht-induziert) über pPha-T1, bzw. Nitrat-induziert über pPha_NR nach Überimpfen und zwei bis fünf tägiger Inkubation transformierter *P. tricornutum* Klone auf Nitrat-haltigem (0,89 mM NO₃⁻) f/2-Festmedium.

6.2.3.10.2 Self-Assembling GFP - *in vivo* Analysen zur Topologie von Membranproteinen

Um die Topologie von putativen Triosephosphat-Translokatoren (TPTs) in den vier Hüllmembranen der komplexen Plastide von *P. tricornutum* zu untersuchen wurde das Self-Assembling GFP System nach Cabantous genutzt (Cabantous et al, 2005). Dieses System beruht auf dem Prinzip der Teilung von GFP in zwei Fragmente (β-Stränge 1-10 und β-Strang 11), welche allein nicht in der Lage sind zu fluoreszieren. Sind beide Fragmente jedoch im gleichen zellulären Kompartiment lokalisiert, kommt es aufgrund sehr hoher Affinität von GFP 1-10 und GFP 11 für das jeweilige andere Fragment zur stabilen Selbstassemblierung und somit zur Rekonstitution der GFP-Fluoreszenz. Das System wurde bereits zuvor erfolgreich zur Analyse der Topologie bzw. für *in vivo* Lokalisationsstudien von Membranproteinen in der komplexen Plastide von *P. tricornutum*, sowie in Apicomplexa genutzt (Bullmann et al, 2010; Cabantous et al, 2005; Hempel et al, 2009; van Dooren et al, 2008).

Jene TPTs die anhand von BTS- und „volle Länge“ (*full length*, FL)-GFP Lokalisationsstudien als Membranproteine der komplexen Plastide von *P. tricornutum* klassifiziert werden konnten, sollten im Weiteren hinsichtlich ihrer spezifischen Topologie innerhalb des Endosymbionten untersucht werden. Dazu wurde GFP 11 als C-terminale Fusion der jeweiligen FL TPT-Sequenzen (1 - 4 und 6) exprimiert. Als Markerproteine für die jeweiligen Subkompartimente der komplexen Plastide, welche C-terminal mit GFP 1-10 fusioniert exprimiert wurden, dienten folgende bereits in vorangehenden Studien erfolgreich getesteten Proteine bzw. Zielsteuerungssequenzen (Bullmann et al, 2010; Hempel et al, 2009): Cytosol, GFP 1-10 ohne Zielsteuerungssequenz; ER, PDI (Protein-Disulfid-Isomerase) in

voller Länge inklusive SP; PPC, BTS von sHsp70; Intermembranraum (IMS), MGD1 (Monogalctosyl Diacylglycerol Synthase) in voller Länge inklusive BTS; Plastidenstroma, BTS von AtpC (Chloroplast ATPase Gamma-Untereinheit). Die jeweiligen Proteinidentifikationsnummern sind im Folgenden aufgelistet: PDI, 44937; sHsp70, 55890; MGD1, 21505; AtpC, 50958. Die Primer zur Amplifikation der jeweiligen Gensequenzen mittels PCR sind Abschnitt 6.1.8 zu entnehmen.

Als Expressionsvektor des Self-Assembling GFP Systems in *P. tricornutum* diene der pPha_Dual_2xNR Vektor, bei dem es sich um ein Derivat von pPha-T1 handelt und der mit zwei *multiple cloning sites* (MCS) ausgestattet ist, welche beide unter der Kontrolle eines induzierbaren endogenen Nitratreduktase-Promotor stehen. In der Regel wurden die mit GFP 1-10 (*BamHI/SacII*) fusionierten Marker-Gensequenzen (*SpeI/BamHI*) über *SpeI/SacII* in die MCSI ligiert, wohingegen die TPT (*EcoRI/BamHI*)-GFP 11 (*BamHI/HindIII*) Fusionssequenzen über *EcoRI/HindIII* in die MCSII von pPha_Dual_2xNR eingebracht wurden. Nach Transformation der Algen unter nicht-induzierenden Bedingungen (Ammonium als einzige Stickstoffquelle) wurden positive Klone in Nitrat-haltiges (0,89 mM NO₃⁻) f/2-Flüssigmedium überführt und für eine optimierte Zeitspanne von 6 h (Hempel et al, 2009) zur Expression der Fusionsproteine *in vivo* induziert, bevor sie unter dem KLSM analysiert werden konnten (6.2.4). Negative Klone, d.h. Klone die keine Fluoreszenz unter dem KLSM zeigten, wurden mittels Kolonie-PCR (6.2.2.4) auf erfolgte genomische Integration der Fusionssequenzen untersucht.

Positiv-, sowie Negativ-Kontrollen für die Anwendung des Self-Assembling GFP Systems in *P. tricornutum* (nicht gezeigt), speziell in der komplexen Plastide, wurden ausgiebig in vorangehenden Arbeiten der AG Maier beschrieben (Bullmann et al, 2010; Hempel et al, 2009).

6.2.3.10.3 Split-GFP - Analysen von *in vivo* Proteininteraktionen

Die Analyse von *in vivo* Protein-Protein-Interaktionen im PPC von *P. tricornutum* erfolgte über die Methode der bi-molekularen Fluoreszenz Komplementierung (BiFC), bzw. das Split-GFP System (Ghosh et al, 2000; Hu et al, 2002; Kerppola, 2006; Park & Raines, 2000). Dabei wird, ähnlich dem Self-Assembling GFP (6.2.3.10.2), GFP in zwei Fragmente gespalten, die einzeln keine Fluoreszenz aufweisen. Der Unterschied zum Self-Assembling GFP ist jedoch, dass die jeweiligen Split-GFP Fragmente (N-GFP und C-GFP) eine ähnliche Größe aufweisen

und nicht spontan assemblieren können wenn diese im gleichen Kompartiment lokalisieren. Lediglich die Interaktion von an N- und C-GFP fusionierten Proteinen kann beide Fragmente in die erforderliche räumliche Nähe bringen und schließlich zur Assemblierung von N- und C-GFP, d.h. zur Rekonstitution der GFP-Fluoreszenz, führen. Das Split-GFP System dient somit dem *in vivo* Nachweis von Protein-Protein-Interaktionen *in situ*, wohingegen das Self-Assembling GFP System Aufschluss über die gemeinsame Lokalisation zweier Proteine bzw. Proteindomänen (bei Membranproteinen) im gleichen Kompartiment gibt.

Die Expression potentieller Interaktionspartner im PPC von *P. tricornutum* fusioniert mit N-, bzw. C-GFP erfolgte über den pPha_Dual_2xNR Vektor. Dazu wurden die Gensequenzen, kodierend für die jeweiligen Proteine von Interesse, 5' and *n-gfp* (*BamHI/SacII*) bzw. *c-gfp* (*BamHI/HindIII*) fusioniert und in die MCSI (*SpeI/SacII*) bzw. MCSII (*EcoRI/HindIII*) von pPha_Dual_2xNR eingebracht. Nach der biolistischen Transformation unter nicht-induzierenden Bedingungen auf f/2-Festmedium (Ammonium als einzige Stickstoffquelle), wurden positive Klone für zwei bis fünf Tage auf Nitrat-haltigem (0,89 mM NO₃⁻) f/2-Festmedium kultiviert um die Expression der Fusionsproteine zu induzieren. Es folgte abschließend die Analyse der *in vivo* Interaktion unter dem KLSM (6.2.4)

Die oben beschriebene Methode des Split-GFP Systems wurde bereits erfolgreich in *P. tricornutum* getestet. Kontrollexperimente, die Spezifität und Funktionalität des Systems in der Diatomee, insbesondere im PPC der komplexen Plastide, validieren (nicht gezeigt), sind in vorangehenden Arbeiten der AG Maier einzusehen (Hempel et al, 2009).

6.2.4 Konfokale Laser Scanning Mikroskopie

Die *in vivo* Analyse exprimierter GFP-Reporterfusionsproteine erfolgte unter dem Konfokalen Laser Scanning Mikroskop TCS SP2 der Firma Leica Microsystems. Es wurden 8 µl einer ggf. Nitrat-induzierten (siehe 6.2.3.9) biolistisch transformierten *P. tricornutum*-Kultur auf einen Objektträger gegeben und unter einem HCX PL APO 40x/1.25-0.75 Oil CS Objektiv fokussiert. Mittels Argon-Laser und Einsatz eines tripeldichroitischen Filters (TD 488/543/633) wurden GFP und Chlorophyll bei 488 nm angeregt. Die Detektion von GFP-Fluoreszenz (500 - 520 nm) und der Autofluoreszenz der Plastide (625 - 720 nm) erfolgte durch Photomultiplier auf unterschiedlichen Kanälen. Die resultierenden Aufnahmen wurden mit der Leica LCS Lite Software analysiert und bearbeitet.

6.2.5 Bioinformatische (*in silico*) Analysen

Die *in silico* Suche nach PPC-lokalisierten Proteinen in *P. tricornutum* wurde auf zwei verschiedene Arten durchgeführt. Der Schwerpunkt lag auf der ersten im Folgenden beschriebenen Vorgehensweise: Um PPC-lokalisierte Proteine, d.h. Proteinsequenzen die eine PPC-spezifische Zielsteuerungssequenz (BTS) besitzen, bioinformatisch zu ermitteln, wurde in einem „gerichteten Ansatz“ nach typischen Faktoren für ein eukaryotes Cytoplasma gesucht. Dies geschah zum einen durch direkte Suche nach annotierten Faktoren in der *P. tricornutum* Genomdatenbank, bzw. NCBI und zum anderen über Sequenzhomologie-Suche mittels BlastP und tBlastN. Für letzteres wurden Proteinsequenzen typisch cytosolischer Proteine aus *Saccharomyces cerevisiae*, *Cyanidioschyzon merolae* oder *Arabidopsis thaliana* als Sequenzvorlagen genutzt (3.1, Tab. 8.1). Im Weiteren wurden die vorhergesagten Genmodelle der ermittelten Proteine auf Vollständigkeit, bzw. auf das Vorhandensein von EST-Daten überprüft. Die authentischen Proteinsequenzen wurden im Anschluss mittels BlastP bzw. NCBI-CDS (Marchler-Bauer et al, 2011) auf die Präsenz konservierter Bereiche und funktioneller Domänen analysiert und dementsprechend eingeordnet. Als nächster Schritt folgte die Untersuchung der Sequenzen nach einer PPC-spezifischen Zielsteuerungssequenz. Dazu wurden die Proteinsequenzen zunächst in den SignalP 3.0 Algorithmus eingespeist. Wurde ein SP mit adäquatem Wert (> 50 %) vorhergesagt, folgte die Analyse einer potentiell folgenden Transitpeptid-ähnlichen Sequenz (TPL) durch TargetP. Da der TargetP-Algorithmus jedoch aufgrund des speziellen Trainingsdatensatzes auf die Transitpeptid-Vorhersage von Proteinen aus Organismen mit primären Plastiden ausgelegt ist, ist die Vorhersage für Transitpeptid-ähnliche Sequenzen sekundär evolvierter Organismen äußerst ungenau. Es wurde daher stattdessen nach einer beliebig langen Extension zwischen putativem SP und konserviertem maturem Protein gesucht, welche für eine Transitpeptid-ähnliche Sequenz in Frage kam und diese wurde schließlich als mögliche PPC-spezifische TPL angenommen. Weitere Kriterien bei der Suche nach potentiellen PPC-Proteinen waren zum einen, dass essentielle cytosolische Proteine mindestens in zwei Kopien vorhanden sein mussten und zum anderen die Detektion ausschließlich als cytosolisch bekannter Faktoren mit vorhergesagtem SP, ohne Rücksicht auf die Länge eines putativen TPs, genügte um als potentiell PPC-lokalisiert eingestuft zu werden.

In einem zweiten „nicht gerichteten“ Ansatz wurden alle Proteinmodelle (10025) aus der *P. tricornutum* Genomdatenbank (*filtered models*, „*best*“, *proteins*, *chromosomes*) extrahiert und mittels SignalP 3.0 analysiert. Diejenigen Proteinmodelle, welche durch den *Hidden*

Markov models SP-Vorhersage-Algorithmus von SignalP 3.0 einen Wert von über 50 % erzielten, wurden als Proteine mit potentielltem SP eingeordnet (2260). In einem weiteren Schritt wurden davon alle potentiell plastidären Proteine, anhand des Vorkommens einer aromatischen Aminosäure oder Leucin an der ersten Position nach der vorhergesagten SP-Schnittstelle (970), aussortiert. Der Großteil der verbleibenden Fraktion (1290 Proteine), sollte demnach aus PPC- (BTS), sowie ER/sekretorischen Proteinen (SP) bestehen. Zumal der Fokus dieser Arbeit auf der Identifizierung löslicher Proteine lag, wurden die 1290 Proteinsequenzen mittels TMHMM auf die Präsenz von potentiellen Transmembrandomänen untersucht (537). Um ER- bzw. sekretorische Proteine von potentiellen PPC-Proteinen zu Unterscheiden, wurden die verbleibenden 753 Proteinsequenzen nach Entfernen der vorhergesagten SP-Sequenz mit TargetP auf die Präsenz eines potentiellen PPC-spezifischen Transitpeptids untersucht. Wie bereits zuvor erwähnt, bestehen enorme Schwächen beim TargetP-Algorithmus hinsichtlich der Transitpeptid-Vorhersage von Organismen mit komplexen Plastiden, weshalb sich dieser Schritt in diesem Ansatz weniger dienlich erwies. Zudem zeigte eine stichprobenartige Analyse der Genmodelle des hier verwendeten Datensatzes aus *P. tricornutum*, dass eine Vielzahl 5' unvollständig vorlag und somit auf Proteinebene, wegen verkürzter oder nicht authentischer N-Termini, keine Klassifizierung dieser Proteinsequenzen anhand genannter Kriterien erfolgen konnte.

Aus diesen Gründen war eine vage Klassifizierung von PPC- und ER/sekretorischen Proteinen aus den 753 Genmodellen nur möglich, indem diese manuell auf Vollständigkeit (EST-Daten) überprüft, mittels BlastP auf konservierte Domänen (potentielle Funktion) untersucht und schließlich auf eine mögliche Extension zwischen SP und konserviertem Bereich (potentielle TPL) analysiert wurden.

7. Literatur

Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup O, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor MF (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J Eukaryot Microbiol* **52**: 399-451

Agrawal S, van Dooren GG, Beatty WL, Striepen B (2009) Genetic evidence that an endosymbiont-derived endoplasmic reticulum-associated protein degradation (ERAD) system functions in import of apicoplast proteins. *J Biol Chem* **284**: 33683-33691

Albiniak AM, Baglieri J, Robinson C (2012) Targeting of luminal proteins across the thylakoid membrane. *J Exp Bot* **63**: 1689-1698

Allen MD, Buchberger A, Bycroft M (2006) The PUB domain functions as a p97 binding module in human peptide N-glycanase. *J Biol Chem* **281**: 25502-25508

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-3402

Anwar K, Klemm RW, Condon A, Severin KN, Zhang M, Ghirlando R, Hu J, Rapoport TA, Prinz WA (2012) The dynamin-like GTPase Sey1p mediates homotypic ER fusion in *S. cerevisiae*. *J Cell Biol* **197**: 209-217

Apt KE, Kroth-Pancic PG, Grossman AR (1996) Stable nuclear transformation of the diatom *Phaeodactylum tricornutum*. *Mol Gen Genet* **252**: 572-579

Apt KE, Zaslavkaia L, Lippmeier JC, Lang M, Kilian O, Wetherbee R, Grossman AR, Kroth PG (2002) In vivo characterization of diatom multipartite plastid targeting signals. *J Cell Sci* **115**: 4061-4069

Archibald JM (2009) The puzzle of plastid evolution. *Curr Biol* **19**: R81-88

Armbrust EV (2009) The life of diatoms in the world's oceans. *Nature* **459**: 185-192

Bachvaroff TR, Sanchez Puerta MV, Delwiche CF (2005) Chlorophyll c-containing plastid relationships based on analyses of a multigene data set with all four chromalveolate lineages. *Mol Biol Evol* **22**: 1772-1782

Badger MR, Price GD (1994) The Role of Carbonic Anhydrase in Photosynthesis. *Annu Rev of Plant Physiol Plant Mol Biol* **45**: 369-392

Balsera M, Soll J, Bolter B (2009) Protein import machineries in endosymbiotic organelles. *Cell Mol Life Sci* **66**: 1903-1923

Barthelme D, Sauer RT (2012) Identification of the Cdc48*20S proteasome as an ancient AAA+ proteolytic machine. *Science* **337**: 843-846

Baugh JM, Viktorova EG, Pilipenko EV (2009) Proteasomes can degrade a significant proportion of cellular proteins independent of ubiquitination. *J Mol Biol* **386**: 814-827

- Baumeister W, Walz J, Zuhl F, Seemuller E (1998) The proteasome: paradigm of a self-compartmentalizing protease. *Cell* **92**: 367-380
- Baurain D, Brinkmann H, Petersen J, Rodriguez-Ezpeleta N, Stechmann A, Demoulin V, Roger AJ, Burger G, Lang BF, Philippe H (2010) Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. *Mol Biol Evol* **27**: 1698-1709
- Bause E (1983) Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. *Biochem J* **209**: 331-336
- Beskow A, Grimberg KB, Bott LC, Salomons FA, Dantuma NP, Young P (2009) A conserved unfoldase activity for the p97 AAA-ATPase in proteasomal degradation. *J Mol Biol* **394**: 732-746
- Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **7**: 1513-1523
- Blankenship RE (1992) Origin and early evolution of photosynthesis. *Photosynth Res* **33**: 91-111
- Blatch GL, Lassle M (1999) The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* **21**: 932-939
- Bodyl A, Stiller JW, Mackiewicz P (2009) Chromalveolate plastids: direct descent or multiple endosymbioses? *Trends Ecol Evol* **24**: 119-121; author reply 121-112
- Bolte K, Bullmann L, Hempel F, Bozarth A, Zauner S, Maier UG (2009) Protein targeting into secondary plastids. *J Eukaryot Microbiol* **56**: 9-15
- Bolte K, Gruenheit N, Felsner G, Sommer MS, Maier UG, Hempel F (2011) Making new out of old: recycling and modification of an ancient protein translocation system during eukaryotic evolution. Mechanistic comparison and phylogenetic analysis of ERAD, SELMA and the peroxisomal importomer. *Bioessays* **33**: 368-376
- Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otiillar RP, Rayko E, Salamov A, Vandepoele K, Beszteri B, Gruber A, Heijde M, Katinka M, Mock T, Valentin K, Verret F, Berges JA, Brownlee C, Cadoret JP, Chiovitti A, Choi CJ, Coesel S, De Martino A, Detter JC, Durkin C, Falciatore A, Fournet J, Haruta M, Huysman MJ, Jenkins BD, Jiroutova K, Jorgensen RE, Joubert Y, Kaplan A, Kroger N, Kroth PG, La Roche J, Lindquist E, Lommer M, Martin-Jezequel V, Lopez PJ, Lucas S, Mangogna M, McGinnis K, Medlin LK, Montsant A, Oudot-Le Secq MP, Napoli C, Obornik M, Parker MS, Petit JL, Porcel BM, Poulsen N, Robison M, Rychlewski L, Ryneerson TA, Schmutz J, Shapiro H, Siaut M, Stanley M, Sussman MR, Taylor AR, Vardi A, von Dassow P, Vyverman W, Willis A, Wyrwicz LS, Rokhsar DS, Weissenbach J, Armbrust EV, Green BR, Van de Peer Y, Grigoriev IV (2008) The Phaeodactylum genome reveals the evolutionary history of diatom genomes. *Nature* **456**: 239-244
- Bowler C, De Martino A, Falciatore A (2010) Diatom cell division in an environmental context. *Curr Opin Plant Biol* **13**: 623-630
- Brands A, Ho TH (2002) Function of a plant stress-induced gene, HVA22. Synthetic enhancement screen with its yeast homolog reveals its role in vesicular traffic. *Plant Physiol* **130**: 1121-1131
- Brooks CF, Johnsen H, van Dooren GG, Muthalagi M, Lin SS, Bohne W, Fischer K, Striepen B (2010) The toxoplasma apicoplast phosphate translocator links cytosolic and apicoplast metabolism and is essential for parasite survival. *Cell Host Microbe* **7**: 62-73
- Bullmann L, Haarmann R, Mirus O, Bredemeier R, Hempel F, Maier UG, Schleiff E (2010) Filling the gap, evolutionarily conserved Omp85 in plastids of chromalveolates. *J Biol Chem* **285**: 6848-6856

- Burki F, Flegontov P, Obornik M, Cihlar J, Pain A, Lukes J, Keeling PJ (2012a) Re-evaluating the green versus red signal in eukaryotes with secondary plastid of red algal origin. *Genome Biol Evol* **4**: 626-635
- Burki F, Okamoto N, Pombert JF, Keeling PJ (2012b) The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. *Proc Biol Sci* **279**: 2246-2254
- Burki F, Shalchian-Tabrizi K, Minge M, Skjaeveland A, Nikolaev SI, Jakobsen KS, Pawlowski J (2007) Phylogenomics reshuffles the eukaryotic supergroups. *PLoS One* **2**: e790
- Cabantous S, Terwilliger TC, Waldo GS (2005) Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat Biotechnol* **23**: 102-107
- Cashmore AR, Jarillo JA, Wu YJ, Liu D (1999) Cryptochromes: blue light receptors for plants and animals. *Science* **284**: 760-765
- Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. *J Eukaryot Microbiol* **46**: 347-366
- Cavalier-Smith T (2000) Membrane heredity and early chloroplast evolution. *Trends Plant Sci* **5**: 174-182
- Cavalier-Smith T (2002) Chloroplast evolution: secondary symbiogenesis and multiple losses. *Curr Biol* **12**: R62-64
- Cavalier-Smith T (2003) Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae). *Philos Trans R Soc Lond B Biol Sci* **358**: 109-133; discussion 133-104
- Chan CX, Yang EC, Banerjee T, Yoon HS, Martone PT, Estevez JM, Bhattacharya D (2011) Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. *Curr Biol* **21**: 328-333
- Coesel S, Mangogna M, Ishikawa T, Heijde M, Rogato A, Finazzi G, Todo T, Bowler C, Falciatore A (2009) Diatom PtCPF1 is a new cryptochrome/photolyase family member with DNA repair and transcription regulation activity. *EMBO Rep* **10**: 655-661
- Coles M, Diercks T, Liermann J, Groger A, Rockel B, Baumeister W, Koretke KK, Lupas A, Peters J, Kessler H (1999) The solution structure of VAT-N reveals a 'missing link' in the evolution of complex enzymes from a simple betaalpha-beta element. *Curr Biol* **9**: 1158-1168
- Colleoni C, Linka M, Deschamps P, Handford MG, Dupree P, Weber AP, Ball SG (2010) Phylogenetic and biochemical evidence supports the recruitment of an ADP-glucose translocator for the export of photosynthate during plastid endosymbiosis. *Mol Biol Evol* **27**: 2691-2701
- Cyr DM, Langer T, Douglas MG (1994) DnaJ-like proteins: molecular chaperones and specific regulators of Hsp70. *Trends Biochem Sci* **19**: 176-181
- Daleke DL (2007) Phospholipid flippases. *J Biol Chem* **282**: 821-825
- De Riso V, Raniello R, Maumus F, Rogato A, Bowler C, Falciatore A (2009) Gene silencing in the marine diatom *Phaeodactylum tricornutum*. *Nucleic Acids Res* **37**: e96

- Deschamps P, Haferkamp I, Dauvillee D, Haebel S, Steup M, Buleon A, Putaux JL, Colleoni C, d'Hulst C, Plancke C, Gould S, Maier U, Neuhaus HE, Ball S (2006) Nature of the periplastidial pathway of starch synthesis in the cryptophyte *Guillardia theta*. *Eukaryot Cell* **5**: 954-963
- Dierks T, Klappa P, Wiech H, Zimmermann R (1993) The role of molecular chaperones in protein transport into the endoplasmic reticulum. *Philos Trans R Soc Lond B Biol Sci* **339**: 335-341
- Dorne AJ, Joyard J, Block MA, Douce R (1985) Localization of phosphatidylcholine in outer envelope membrane of spinach chloroplasts. *J Cell Biol* **100**: 1690-1697
- Dorrell RG, Smith AG (2011) Do red and green make brown?: perspectives on plastid acquisitions within chromalveolates. *Eukaryot Cell* **10**: 856-868
- Douglas S, Zauner S, Fraunholz M, Beaton M, Penny S, Deng LT, Wu X, Reith M, Cavalier-Smith T, Maier UG (2001) The highly reduced genome of an enslaved algal nucleus. *Nature* **410**: 1091-1096
- Douglas SE, Turner S (1991) Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor. *J Mol Evol* **33**: 267-273
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* **12**: 13-15
- Dreveny I, Pye VE, Beuron F, Briggs LC, Isaacson RL, Matthews SJ, McKeown C, Yuan X, Zhang X, Freemont PS (2004) p97 and close encounters of every kind: a brief review. *Biochem Soc Trans* **32**: 715-720
- Duy D, Soll J, Philippar K (2007) Solute channels of the outer membrane: from bacteria to chloroplasts. *Biol Chem* **388**: 879-889
- Farmer LM, Book AJ, Lee KH, Lin YL, Fu H, Vierstra RD (2010) The RAD23 family provides an essential connection between the 26S proteasome and ubiquitylated proteins in Arabidopsis. *Plant Cell* **22**: 124-142
- Fast NM, Kissinger JC, Roos DS, Keeling PJ (2001) Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. *Mol Biol Evol* **18**: 418-426
- Felsner G, Sommer MS, Gruenheit N, Hempel F, Moog D, Zauner S, Martin W, Maier UG (2011) ERAD components in organisms with complex red plastids suggest recruitment of a preexisting protein transport pathway for the periplastid membrane. *Genome Biol Evol* **3**: 140-150
- Felsner G, Sommer MS, Maier UG (2010) The physical and functional borders of transit peptide-like sequences in secondary endosymbionts. *BMC Plant Biol* **10**: 223
- Fleige T, Fischer K, Ferguson DJ, Gross U, Bohne W (2007) Carbohydrate metabolism in the *Toxoplasma gondii* apicoplast: localization of three glycolytic isoenzymes, the single pyruvate dehydrogenase complex, and a plastid phosphate translocator. *Eukaryot Cell* **6**: 984-996
- Flügge UI (1999) Phosphate Translocators in Plastids. *Annu Rev Plant Physiol Plant Mol Biol* **50**: 27-45
- Frasch AC (2000) Functional diversity in the trans-sialidase and mucin families in *Trypanosoma cruzi*. *Parasitol Today* **16**: 282-286
- Gallastegui N, Groll M (2010) The 26S proteasome: assembly and function of a destructive machine. *Trends Biochem Sci* **35**: 634-642

- Galway ME, Heckman JW, Jr., Schiefelbein JW (1997) Growth and ultrastructure of Arabidopsis root hairs: the *rhd3* mutation alters vacuole enlargement and tip growth. *Planta* **201**: 209-218
- Gao H, Kadirjan-Kalbach D, Froehlich JE, Osteryoung KW (2003) ARC5, a cytosolic dynamin-like protein from plants, is part of the chloroplast division machinery. *Proc Natl Acad Sci U S A* **100**: 4328-4333
- Ghosh I, Hamilton AD, Regan L (2000) Antiparallel leucine zipperdirected protein reassembly: application to the green fluorescent protein. *J Am Chem Soc* **122**: 5658-5659
- Gibbs SP (1979) The route of entry of cytoplasmically synthesized proteins into chloroplasts of algae possessing chloroplast ER. *J Cell Sci* **35**: 253-266
- Gilson PR (2001) Nucleomorph genomes: much ado about practically nothing. *Genome Biol* **2**: REVIEWS1022
- Gorres KL, Raines RT (2010) Prolyl 4-hydroxylase. *Crit Rev Biochem Mol Biol* **45**: 106-124
- Gould SB, Sommer MS, Hadfi K, Zauner S, Kroth PG, Maier UG (2006a) Protein targeting into the complex plastid of cryptophytes. *J Mol Evol* **62**: 674-681
- Gould SB, Sommer MS, Kroth PG, Gile GH, Keeling PJ, Maier UG (2006b) Nucleus-to-nucleus gene transfer and protein retargeting into a remnant cytoplasm of cryptophytes and diatoms. *Mol Biol Evol* **23**: 2413-2422
- Gould SB, Waller RF, McFadden GI (2008) Plastid evolution. *Annu Rev Plant Biol* **59**: 491-517
- Grant CM (2001) Role of the glutathione/glutaredoxin and thioredoxin systems in yeast growth and response to stress conditions. *Mol Microbiol* **39**: 533-541
- Green BR (2011) After the primary endosymbiosis: an update on the chromalveolate hypothesis and the origins of algae with Chl c. *Photosynth Res* **107**: 103-115
- Griac P (2007) Sec14 related proteins in yeast. *Biochim Biophys Acta* **1771**: 737-745
- Groll M, Heinemeyer W, Jager S, Ullrich T, Bochtler M, Wolf DH, Huber R (1999) The catalytic sites of 20S proteasomes and their role in subunit maturation: a mutational and crystallographic study. *Proc Natl Acad Sci U S A* **96**: 10976-10983
- Gruber A, Vugrinec S, Hempel F, Gould SB, Maier UG, Kroth PG (2007) Protein targeting into complex diatom plastids: functional characterisation of a specific targeting motif. *Plant Mol Biol* **64**: 519-530
- Gruber A, Weber T, Bartulos CR, Vugrinec S, Kroth PG (2009) Intracellular distribution of the reductive and oxidative pentose phosphate pathways in two diatoms. *J Basic Microbiol* **49**: 58-72
- Grune T, Merker K, Sandig G, Davies KJ (2003) Selective degradation of oxidatively modified protein substrates by the proteasome. *Biochem Biophys Res Commun* **305**: 709-718
- Hackett JD, Yoon HS, Li S, Reyes-Prieto A, Rummele SE, Bhattacharya D (2007) Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates. *Mol Biol Evol* **24**: 1702-1713
- Haferkamp I, Deschamps P, Ast M, Jeblick W, Maier U, Ball S, Neuhaus HE (2006) Molecular and biochemical analysis of periplastidial starch metabolism in the cryptophyte *Guillardia theta*. *Eukaryot Cell* **5**: 964-971

- Harper JT, Keeling PJ (2003) Nucleus-encoded, plastid-targeted glyceraldehyde-3-phosphate dehydrogenase (GAPDH) indicates a single origin for chromalveolate plastids. *Mol Biol Evol* **20**: 1730-1735
- Harper JT, Waanders E, Keeling PJ (2005) On the monophyly of chromalveolates using a six-protein phylogeny of eukaryotes. *Int J Syst Evol Microbiol* **55**: 487-496
- Hartl FU, Hayer-Hartl M (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**: 1852-1858
- Harvey SH, Krien MJ, O'Connell MJ (2002) Structural maintenance of chromosomes (SMC) proteins, a family of conserved ATPases. *Genome Biol* **3**: REVIEWS3003
- Hashimoto H (2005) The ultrastructural features and division of secondary plastids. *J Plant Res* **118**: 163-172
- Heintzelman MB, Enriquez ME (2010) Myosin diversity in the diatom *Phaeodactylum tricornutum*. *Cytoskeleton (Hoboken)* **67**: 142-151
- Hempel F, Bozarth AS, Lindenkamp N, Klingl A, Zauner S, Linne U, Steinbuchel A, Maier UG (2011) Microalgae as bioreactors for bioplastic production. *Microb Cell Fact* **10**: 81
- Hempel F, Bullmann L, Lau J, Zauner S, Maier UG (2009) ERAD-derived preprotein transport across the second outermost plastid membrane of diatoms. *Mol Biol Evol* **26**: 1781-1790
- Hempel F, Felsner G, Maier UG (2010) New mechanistic insights into pre-protein transport across the second outermost plastid membrane of diatoms. *Mol Microbiol* **76**: 793-801
- Holmgren A (1985) Thioredoxin. *Annu Rev Biochem* **54**: 237-271
- Hopkinson BM, Dupont CL, Allen AE, Morel FM (2011) Efficiency of the CO₂-concentrating mechanism of diatoms. *Proc Natl Acad Sci U S A* **108**: 3830-3837
- Hoyt MA, Coffino P (2004) Ubiquitin-free routes into the proteasome. *Cell Mol Life Sci* **61**: 1596-1600
- Hu CD, Chinenov Y, Kerppola TK (2002) Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol Cell* **9**: 789-798
- Issoglio FM, Carrizo ME, Romero JM, Curtino JA (2012) Mechanisms of monomeric and dimeric glycogenin autoglucosylation. *J Biol Chem* **287**: 1955-1961
- Janouskovec J, Horak A, Obornik M, Lukes J, Keeling PJ (2010) A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proc Natl Acad Sci U S A* **107**: 10949-10954
- Jarvis P (2008) Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytol* **179**: 257-285
- Joyard J, Teyssier E, Mieg C, Berny-Seigneurin D, Marechal E, Block MA, Dorne AJ, Rolland N, Ajlani G, Douce R (1998) The biochemical machinery of plastid envelope membranes. *Plant Physiol* **118**: 715-723
- Kalanon M, McFadden GI (2008) The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*: a bioinformatic comparison of Toc and Tic components in plants, green algae and red algae. *Genetics* **179**: 95-112

- Kalanon M, Tonkin CJ, McFadden GI (2009) Characterization of two putative protein translocation components in the apicoplast of *Plasmodium falciparum*. *Eukaryot Cell* **8**: 1146-1154
- Kaplan A, Reinhold L (1999) Co₂ Concentrating Mechanisms in Photosynthetic Microorganisms. *Annu Rev Plant Physiol Plant Mol Biol* **50**: 539-570
- Keeling PJ (2009) Chromalveolates and the evolution of plastids by secondary endosymbiosis. *J Eukaryot Microbiol* **56**: 1-8
- Keeling PJ (2010) The endosymbiotic origin, diversification and fate of plastids. *Philos Trans R Soc Lond B Biol Sci* **365**: 729-748
- Keeling PJ, Deane JA, Hink-Schauer C, Douglas SE, Maier UG, McFadden GI (1999) The secondary endosymbiont of the cryptomonad *Guillardia theta* contains alpha-, beta-, and gamma-tubulin genes. *Mol Biol Evol* **16**: 1308-1313
- Kerppola TK (2006) Visualization of molecular interactions by fluorescence complementation. *Nat Rev Mol Cell Biol* **7**: 449-456
- Kilian O, Kroth PG (2004) Presequence acquisition during secondary endocytobiosis and the possible role of introns. *J Mol Evol* **58**: 712-721
- Kilian O, Kroth PG (2005) Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids. *Plant J* **41**: 175-183
- Kitao Y, Harada H, Matsuda Y (2008) Localization and targeting mechanisms of two chloroplastic beta-carbonic anhydrases in the marine diatom *Phaeodactylum tricornutum*. *Physiol Plant* **133**: 68-77
- Kleine T, Maier UG, Leister D (2009) DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu Rev Plant Biol* **60**: 115-138
- Kordeli E, Lambert S, Bennett V (1995) AnkyrinG. A new ankyrin gene with neural-specific isoforms localized at the axonal initial segment and node of Ranvier. *J Biol Chem* **270**: 2352-2359
- Kraut J (1977) Serine proteases: structure and mechanism of catalysis. *Annu Rev Biochem* **46**: 331-358
- Kroth PG, Chiovitti A, Gruber A, Martin-Jezequel V, Mock T, Parker MS, Stanley MS, Kaplan A, Caron L, Weber T, Maheswari U, Armbrust EV, Bowler C (2008) A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. *PLoS One* **3**: e1426
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685
- Lane CE, van den Heuvel K, Kozera C, Curtis BA, Parsons BJ, Bowman S, Archibald JM (2007) Nucleomorph genome of *Hemiselmis andersenii* reveals complete intron loss and compaction as a driver of protein structure and function. *Proc Natl Acad Sci U S A* **104**: 19908-19913
- Lang M, Apt KE, Kroth PG (1998) Protein transport into "complex" diatom plastids utilizes two different targeting signals. *J Biol Chem* **273**: 30973-30978
- Latterich M, Frohlich KU, Schekman R (1995) Membrane fusion and the cell cycle: Cdc48p participates in the fusion of ER membranes. *Cell* **82**: 885-893

- Laufen T, Mayer MP, Beisel C, Klostermeier D, Mogk A, Reinstein J, Bukau B (1999) Mechanism of regulation of hsp70 chaperones by DnaJ cochaperones. *Proc Natl Acad Sci U S A* **96**: 5452-5457
- Le Corguille G, Pearson G, Valente M, Viegas C, Gschloessl B, Corre E, Bailly X, Peters AF, Jubin C, Vacherie B, Cock JM, Leblanc C (2009) Plastid genomes of two brown algae, *Ectocarpus siliculosus* and *Fucus vesiculosus*: further insights on the evolution of red-algal derived plastids. *BMC Evol Biol* **9**: 253
- Lev S (2010) Non-vesicular lipid transport by lipid-transfer proteins and beyond. *Nat Rev Mol Cell Biol* **11**: 739-750
- Lim L, Linka M, Mullin KA, Weber AP, McFadden GI (2010) The carbon and energy sources of the non-photosynthetic plastid in the malaria parasite. *FEBS Lett* **584**: 549-554
- Lim L, McFadden GI (2010) The evolution, metabolism and functions of the apicoplast. *Philos Trans R Soc Lond B Biol Sci* **365**: 749-763
- Linka M, Jamai A, Weber AP (2008) Functional characterization of the plastidic phosphate translocator gene family from the thermo-acidophilic red alga *Galdieria sulphuraria* reveals specific adaptations of primary carbon partitioning in green plants and red algae. *Plant Physiol* **148**: 1487-1496
- Lomako J, Lomako WM, Whelan WJ (2004) Glycogenin: the primer for mammalian and yeast glycogen synthesis. *Biochim Biophys Acta* **1673**: 45-55
- Lowe J, Stock D, Jap B, Zwickl P, Baumeister W, Huber R (1995) Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science* **268**: 533-539
- Madsen L, Seeger M, Semple CA, Hartmann-Petersen R (2009) New ATPase regulators--p97 goes to the PUB. *Int J Biochem Cell Biol* **41**: 2380-2388
- Maier UG, Douglas SE, Cavalier-Smith T (2000) The nucleomorph genomes of cryptophytes and chlorarachniophytes. *Protist* **151**: 103-109
- Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* **39**: D225-229
- Maréchal E, Block MA, Dorne AJ, Douce R, Joyard J (1997) Lipid synthesis and metabolism in the plastid envelope. *Physiol Plant* **100**: 65-77
- Margulis L (1970) Origin of eucaryotic cells. *Yale University Press, New Haven*
- Marks B, Stowell MH, Vallis Y, Mills IG, Gibson A, Hopkins CR, McMahon HT (2001) GTPase activity of dynamin and resulting conformation change are essential for endocytosis. *Nature* **410**: 231-235
- Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc Natl Acad Sci U S A* **99**: 12246-12251
- May T, Soll J (2000) 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. *Plant Cell* **12**: 53-64

- McAndrew RS, Froehlich JE, Vitha S, Stokes KD, Osteryoung KW (2001) Colocalization of plastid division proteins in the chloroplast stromal compartment establishes a new functional relationship between FtsZ1 and FtsZ2 in higher plants. *Plant Physiol* **127**: 1656-1666
- McFadden GI, van Dooren GG (2004) Evolution: red algal genome affirms a common origin of all plastids. *Curr Biol* **14**: R514-516
- Medalia N, Beer A, Zwickl P, Mihalache O, Beck M, Medalia O, Navon A (2009) Architecture and molecular mechanism of PAN, the archaeal proteasome regulatory ATPase. *J Biol Chem* **284**: 22952-22960
- Mereschkowsky C (1905) Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol Zentralbl* **25**: 593-604
- Mernberger M, Moog D, Stork S, Zauner S, Maier U, Hüllermeier E (2011) Prediction of protein localization for specialized compartments using time series kernels. *Proc GCB, LNI (conference paper)*
- Meusser B, Hirsch C, Jarosch E, Sommer T (2005) ERAD: the long road to destruction. *Nat Cell Biol* **7**: 766-772
- Meyer H, Bug M, Bremer S (2012) Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system. *Nat Cell Biol* **14**: 117-123
- Micheli F (2001) Pectin methylesterases: cell wall enzymes with important roles in plant physiology. *Trends Plant Sci* **6**: 414-419
- Miyagishima S, Kuroiwa H, Kuroiwa T (2001) The timing and manner of disassembly of the apparatuses for chloroplast and mitochondrial division in the red alga *Cyanidioschyzon merolae*. *Planta* **212**: 517-528
- Miyagishima SY, Kabeya Y (2010) Chloroplast division: squeezing the photosynthetic captive. *Curr Opin Microbiol* **13**: 738-746
- Miyagishima SY, Nishida K, Mori T, Matsuzaki M, Higashiyama T, Kuroiwa H, Kuroiwa T (2003) A plant-specific dynamin-related protein forms a ring at the chloroplast division site. *Plant Cell* **15**: 655-665
- Moog D (2009) Heterologe Lokalisationsstudien symbiontspezifischer ERAD-Faktoren aus *Emiliania huxleyi* in *Phaeodactylum tricornutum* und PPC-Proteomik in *Phaeodactylum tricornutum*: Analyse eines putativ Vesikel-vermittelten Proteintransports und Identifizierung PPC-lokalisierter Faktoren. Philipps-Universität Marburg, Masterarbeit,
- Moog D, Stork S, Zauner S, Maier UG (2011) In silico and in vivo investigations of proteins of a minimized eukaryotic cytoplasm. *Genome Biol Evol* **3**: 375-382
- Moore CE, Archibald JM (2009) Nucleomorph genomes. *Annu Rev Genet* **43**: 251-264
- Moore CE, Curtis B, Mills T, Tanifuji G, Archibald JM (2012) Nucleomorph Genome Sequence of the Cryptophyte Alga *Chroomonas mesostigmatica* CCMP1168 Reveals Lineage-Specific Gene Loss and Genome Complexity. *Genome Biol Evol* **4**: 1162-1175
- Mosavi LK, Cammett TJ, Desrosiers DC, Peng ZY (2004) The ankyrin repeat as molecular architecture for protein recognition. *Protein Sci* **13**: 1435-1448

- Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D (2009) Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* **324**: 1724-1726
- Mullin KA, Lim L, Ralph SA, Spurck TP, Handman E, McFadden GI (2006) Membrane transporters in the relict plastid of malaria parasites. *Proc Natl Acad Sci U S A* **103**: 9572-9577
- Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* **51 Pt 1**: 263-273
- Murata S, Yashiroda H, Tanaka K (2009) Molecular mechanisms of proteasome assembly. *Nat Rev Mol Cell Biol* **10**: 104-115
- Nicholas KB, Nicholas HB (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments. <http://www.wpscedu/biomed/genedoc>
- Orlowski M, Wilk S (2003) Ubiquitin-independent proteolytic functions of the proteasome. *Arch Biochem Biophys* **415**: 1-5
- Park SH, Raines RT (2000) Green fluorescent protein chimeras to probe protein-protein interactions. *Methods Enzymol* **328**: 251-261
- Patron NJ, Inagaki Y, Keeling PJ (2007) Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. *Curr Biol* **17**: 887-891
- Patron NJ, Rogers MB, Keeling PJ (2004) Gene replacement of fructose-1,6-bisphosphate aldolase supports the hypothesis of a single photosynthetic ancestor of chromalveolates. *Eukaryot Cell* **3**: 1169-1175
- Patron NJ, Waller RF (2007) Transit peptide diversity and divergence: A global analysis of plastid targeting signals. *Bioessays* **29**: 1048-1058
- Patron NJ, Waller RF, Archibald JM, Keeling PJ (2005) Complex protein targeting to dinoflagellate plastids. *J Mol Biol* **348**: 1015-1024
- Perona JJ, Craik CS (1995) Structural basis of substrate specificity in the serine proteases. *Protein Sci* **4**: 337-360
- Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber AP, Schwacke R, Gross J, Blouin NA, Lane C, Reyes-Prieto A, Durnford DG, Neilson JA, Lang BF, Burger G, Steiner JM, Löffelhardt W, Meuser JE, Posewitz MC, Ball S, Arias MC, Henrissat B, Coutinho PM, Rensing SA, Symeonidi A, Doddapaneni H, Green BR, Rajah VD, Boore J, Bhattacharya D (2012) Cyanophora paradoxa genome elucidates origin of photosynthesis in algae and plants. *Science* **335**: 843-847
- Qiu XB, Shao YM, Miao S, Wang L (2006) The diversity of the DnaJ/Hsp40 family, the crucial partners for Hsp70 chaperones. *Cell Mol Life Sci* **63**: 2560-2570
- Radhamony RN, Theg SM (2006) Evidence for an ER to Golgi to chloroplast protein transport pathway. *Trends Cell Biol* **16**: 385-387
- Reumann S, Inoue K, Keegstra K (2005) Evolution of the general protein import pathway of plastids (review). *Mol Membr Biol* **22**: 73-86
- Reyes-Prieto A, Weber AP, Bhattacharya D (2007) The origin and establishment of the plastid in algae and plants. *Annu Rev Genet* **41**: 147-168

- Roach PJ, Depaoli-Roach AA, Hurley TD, Tagliabracci VS (2012) Glycogen and its metabolism: some new developments and old themes. *Biochem J* **441**: 763-787
- Rodriguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Löffelhardt W, Bohnert HJ, Philippe H, Lang BF (2005) Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr Biol* **15**: 1325-1330
- Rogers MB, Gilson PR, Su V, McFadden GI, Keeling PJ (2007) The complete chloroplast genome of the chlorarachniophyte *Bigeloviella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. *Mol Biol Evol* **24**: 54-62
- Rubin CS, Rosen OM (1975) Protein phosphorylation. *Annu Rev Biochem* **44**: 831-887
- Saito K, Tautz L, Mustelin T (2007) The lipid-binding SEC14 domain. *Biochim Biophys Acta* **1771**: 719-726
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press, New York*
- Sanchez Puerta MV, Delwiche CF (2008) A hypothesis for plastid evolution in chromalveolates. *J Phycol* **44**: 1097-1107
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**: 5463-5467
- Schaffner W, Weissmann C (1973) A rapid, sensitive, and specific method for the determination of protein in dilute solution. *Anal Biochem* **56**: 502-514
- Schlegel T, Mirus O, von Haeseler A, Schleiff E (2007) The tetratricopeptide repeats of receptors involved in protein translocation across membranes. *Mol Biol Evol* **24**: 2763-2774
- Schnabl M, Oskolkova OV, Holic R, Brezna B, Pichler H, Zagorsek M, Kohlwein SD, Paltauf F, Daum G, Griac P (2003) Subcellular localization of yeast Sec14 homologues and their involvement in regulation of phospholipid turnover. *Eur J Biochem* **270**: 3133-3145
- Sedgwick SG, Smerdon SJ (1999) The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biochem Sci* **24**: 311-316
- Sheiner L, Demerly JL, Poulsen N, Beatty WL, Lucas O, Behnke MS, White MW, Striepen B (2011) A systematic screen to discover and analyze apicoplast proteins identifies a conserved and essential protein import factor. *PLoS Pathog* **7**: e1002392
- Sheiner L, Striepen B (2012) Protein sorting in complex plastids. *Biochim Biophys Acta*
- Shi LX, Theg SM (2012) The chloroplast protein import system: From algae to trees. *Biochim Biophys Acta*
- Soll J, Schleiff E (2004) Protein import into chloroplasts. *Nat Rev Mol Cell Biol* **5**: 198-208
- Sommer M, Rudolf M, Tillmann B, Tripp J, Sommer MS, Schleiff E (2012) Toc33 and Toc64-III cooperate in precursor protein import into the chloroplasts of *Arabidopsis thaliana*. *Plant Cell Environ*
- Sommer MS, Gould SB, Lehmann P, Gruber A, Przyborski JM, Maier UG (2007) Der1-mediated preprotein import into the periplastid compartment of chromalveolates? *Mol Biol Evol* **24**: 918-928

- Spork S, Hiss JA, Mandel K, Sommer M, Kooij TW, Chu T, Schneider G, Maier UG, Przyborski JM (2009) An unusual ERAD-like complex is targeted to the apicoplast of *Plasmodium falciparum*. *Eukaryot Cell* **8**: 1134-1145
- Stefano G, Renna L, Moss T, McNew JA, Brandizzi F (2012) In *Arabidopsis*, the spatial and dynamic organization of the endoplasmic reticulum and Golgi apparatus is influenced by the integrity of the C-terminal domain of RHD3, a non-essential GTPase. *Plant J* **69**: 957-966
- Steiner JM, Yusa F, Pompe JA, Löffelhardt W (2005) Homologous protein import machineries in chloroplasts and cyanelles. *Plant J* **44**: 646-652
- Stoebe B, Maier UG (2002) One, two, three: nature's tool box for building plastids. *Protoplasma* **219**: 123-130
- Stolz A, Hilt W, Buchberger A, Wolf DH (2011) Cdc48: a power machine in protein degradation. *Trends Biochem Sci* **36**: 515-523
- Stork S, Moog D, Przyborski JM, Wilhelmi I, Zauner S, Maier UG (2012) Distribution of the SELMA translocon in secondary plastids of red algal origin and predicted uncoupling of ubiquitin-dependent translocation from degradation. *Eukaryot Cell*: doi:10.1128/EC.00183-00112
- Strittmatter P, Soll J, Bolter B (2010) The chloroplast protein import machinery: a review. *Methods Mol Biol* **619**: 307-321
- Strunnikov AV, Jessberger R (1999) Structural maintenance of chromosomes (SMC) proteins: conserved molecular properties for multiple biological functions. *Eur J Biochem* **263**: 6-13
- Suzuki T, Park H, Till EA, Lennarz WJ (2001) The PUB domain: a putative protein-protein interaction domain implicated in the ubiquitin-proteasome pathway. *Biochem Biophys Res Commun* **287**: 1083-1087
- Tachibana M, Allen AE, Kikutani S, Endo Y, Bowler C, Matsuda Y (2011) Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. *Photosynth Res* **109**: 205-221
- Tanaka Y, Nakatsuma D, Harada H, Ishida M, Matsuda Y (2005) Localization of soluble beta-carbonic anhydrase in the marine diatom *Phaeodactylum tricornutum*. Sorting to the chloroplast and cluster formation on the girdle lamellae. *Plant Physiol* **138**: 207-217
- Tanifuji G, Onodera NT, Wheeler TJ, Dlutek M, Donaher N, Archibald JM (2011) Complete nucleomorph genome sequence of the nonphotosynthetic alga *Cryptomonas paramecium* reveals a core nucleomorph gene set. *Genome Biol Evol* **3**: 44-54
- Tarentino AL, Plummer TH, Jr. (1994) Enzymatic deglycosylation of asparagine-linked glycans: purification, properties, and specificity of oligosaccharide-cleaving enzymes from *Flavobacterium meningosepticum*. *Methods Enzymol* **230**: 44-57
- Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics* **Chapter 2**: Unit 2 3
- Tomko RJ, Jr., Hochstrasser M (2011) Order of the proteasomal ATPases and eukaryotic proteasome assembly. *Cell Biochem Biophys* **60**: 13-20
- Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**: 4350-4354

- van Dooren GG, Reiff SB, Tomova C, Meissner M, Humbel BM, Striepen B (2009) A novel dynamin-related protein has been recruited for apicoplast fission in *Toxoplasma gondii*. *Curr Biol* **19**: 267-276
- van Dooren GG, Tomova C, Agrawal S, Humbel BM, Striepen B (2008) *Toxoplasma gondii* Tic20 is essential for apicoplast protein import. *Proc Natl Acad Sci U S A* **105**: 13574-13579
- Vardi A, Thamtrakoln K, Bidle KD, Falkowski PG (2008) Diatom genomes come of age. *Genome Biol* **9**: 245
- Villarejo A, Buren S, Larsson S, Dejardin A, Monne M, Rudhe C, Karlsson J, Jansson S, Lerouge P, Rolland N, von Heijne G, Grebe M, Bako L, Samuelsson G (2005) Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nat Cell Biol* **7**: 1224-1231
- Voges D, Zwickl P, Baumeister W (1999) The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu Rev Biochem* **68**: 1015-1068
- Wang Q, Song C, Li CC (2003) Hexamerization of p97-VCP is promoted by ATP binding to the D1 domain and required for ATPase and biological activities. *Biochem Biophys Res Commun* **300**: 253-260
- Weber AP (2004) Solute transporters as connecting elements between cytosol and plastid stroma. *Curr Opin Plant Biol* **7**: 247-253
- Weber AP, Linka M, Bhattacharya D (2006) Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor. *Eukaryot Cell* **5**: 609-612
- Weber AP, Linka N (2011) Connecting the plastid: transporters of the plastid envelope and their role in linking plastidial with cytosolic metabolism. *Annu Rev Plant Biol* **62**: 53-77
- Weber T, Gruber A, Kroth PG (2009) The presence and localization of thioredoxins in diatoms, unicellular algae of secondary endosymbiotic origin. *Mol Plant* **2**: 468-477
- Wickstead B, Gull K (2011) The evolution of the cytoskeleton. *J Cell Biol* **194**: 513-525
- Wilkinson B, Gilbert HF (2004) Protein disulfide isomerase. *Biochim Biophys Acta* **1699**: 35-44
- Woehle C, Dagan T, Martin WF, Gould SB (2011) Red and problematic green phylogenetic signals among thousands of nuclear genes from the photosynthetic and apicomplexa-related *Chromera velia*. *Genome Biol Evol* **3**: 1220-1230
- Wolf DH, Stolz A (2011) The Cdc48 machine in endoplasmic reticulum associated protein degradation. *Biochim Biophys Acta* **1823**: 117-124
- Yang Y, Glynn JM, Olson BJ, Schmitz AJ, Osteryoung KW (2008) Plastid division: across time and space. *Curr Opin Plant Biol* **11**: 577-584
- Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D (2004) A molecular timeline for the origin of photosynthetic eukaryotes. *Mol Biol Evol* **21**: 809-818
- Yoon HS, Hackett JD, Pinto G, Bhattacharya D (2002) The single, ancient origin of chromist plastids. *Proc Natl Acad Sci U S A* **99**: 15507-15512
- Yoshida Y, Kuroiwa H, Misumi O, Nishida K, Yagisawa F, Fujiwara T, Nanamiya H, Kawamura F, Kuroiwa T (2006) Isolated chloroplast division machinery can actively constrict after stretching. *Science* **313**: 1435-1438

Yoshida Y, Kuroiwa H, Misumi O, Yoshida M, Ohnuma M, Fujiwara T, Yagisawa F, Hirooka S, Imoto Y, Matsushita K, Kawano S, Kuroiwa T (2010) Chloroplasts divide by contraction of a bundle of nanofilaments consisting of polyglucan. *Science* **329**: 949-953

Yoshida Y, Miyagishima SY, Kuroiwa H, Kuroiwa T (2012) The plastid-dividing machinery: formation, constriction and fission. *Curr Opin Plant Biol*

Young JC, Barral JM, Ulrich Hartl F (2003) More than folding: localized functions of cytosolic chaperones. *Trends Biochem Sci* **28**: 541-547

Zaslavskaja LA, Lippmeier JC, Kroth PG, Grossman AR, Apt KE (2000) Transformation of the diatom *Phaeodactylum tricornutum* (Bacillariophyceae) with a variety of selectable marker and reporter genes. *J Phycol* **36**: 379-386

Zauner S, Fraunholz M, Wastl J, Penny S, Beaton M, Cavalier-Smith T, Maier UG, Douglas S (2000) Chloroplast protein and centrosomal genes, a tRNA intron, and odd telomeres in an unusually compact eukaryotic genome, the cryptomonad nucleomorph. *Proc Natl Acad Sci U S A* **97**: 200-205

Zhao G, Zhou X, Wang L, Li G, Schindelin H, Lennarz WJ (2007) Studies on peptide:N-glycanase-p97 interaction suggest that p97 phosphorylation modulates endoplasmic reticulum-associated degradation. *Proc Natl Acad Sci U S A* **104**: 8785-8790

8. Anhang

8.1. Tabellarische Zusammenfassung der Ergebnisse der kombinierten *in silico/in vivo* Ansätze zur Identifizierung PPC-lokalisierter Proteine in *P. tricornutum*

Die im Folgenden aufgeführten Proteine wurden, wie unter 6.2.5 beschrieben, entweder über den gerichteten Ansatz durch BlastP bzw. direkte Suche nach Annotation, oder den nicht gerichteten Ansatz (Zeilen grau hinterlegt) in der *P. tricornutum* Genomdatenbank und NCBI Proteindatenbank identifiziert. Alle Proteinmodelle wurden mittels SignalP, TargetP und NCBI BlastP (Detektion konservierter Bereiche bzw. N-terminaler Extensionen) auf die Präsenz N-terminaler Zielsteuerungssequenzen untersucht. Dazu wurden zunächst alle Genmodelle hinsichtlich vorhandener EST-Daten überprüft, um zu gewährleisten, dass die Zielsteuerungssequenz-Vorhersage mit den authentischen N-Termini der Proteine durchgeführt werden konnte. Bei Genmodellen für die 5' keine EST-Daten vorlagen, wurde das erste mögliche Startcodon für die Lokalisationsvorhersage der kodierten Proteine verwendet. Proteine für die durch SignalP und TargetP keine N-terminale Zielsteuerungssequenz ermittelt werden konnte wurden als cytosolisch eingestuft. Es ist jedoch nicht auszuschließen, dass diese Proteine möglicherweise aufgrund der Präsenz anderer, nicht N-terminaler Zielsteuerungssequenzen auch in anderen Kompartimenten lokalisiert sein könnten.

Tab. 8.1 Ergebnisse der kombinierten *in silico/in vivo* Ansätze zur Identifizierung PPC-lokalisierter Proteine in *P. tricornutum*

Prozess	ID	Annotierung / Bezeichnung	Putative Funktion / Domänen	(Putative) Lokalisation	Genmodell
ERAD/SELMA	54754 ²	hUba1	Ubiquitin-activating enzyme E1	C	DNA/EST
	54460 ²	sUba1	Ubiquitin-activating enzyme E1	PPC	DNA
	34112 ⁵	sUbc4	Ubiquitin-conjugating enzyme E2	PPC	DNA
	37077 ²	hUbc6	Ubiquitin-conjugating enzyme E2	C	DNA
	56431 ²	sUbc6	Ubiquitin-conjugating enzyme E2	PPC	DNA
	44372 ²	hUfd1	Ubiquitin fusion degradation	C	DNA/EST

Proteasomale Untereinheiten	49319 ²	sUfd1	Ubiquitin fusion degradation	PPC	DNA/EST
	56440 ²	hUb	Ubiquitin	C	DNA/EST
	54323 ²	sUb	Ubiquitin	PPC	DNA/EST
	46966 ^{2,3}	hDer1-1	Derlin, Degradation at the ER, rhomboid family member	ER (MP)	DNA/EST
	31697 ^{2,3}	sDer1-1	Derlin, Degradation at the ER, rhomboid family member	PPC (MP)	DNA/EST
	37614 ^{2,3}	hDer1-2	Derlin, Degradation at the ER, rhomboid family member	ER (MP)	DNA/EST
	35965 ^{2,3}	sDer1-2	Derlin, Degradation at the ER, rhomboid family member	PPC (MP)	DNA/EST
	49801 ²	hHrd1	Ubiquitin-protein ligase E3, HMG-CoA reductase degradation protein 1	ER (MP)	DNA/EST
	48035 ⁴	ptE3P	<i>P. tricornutum</i> E3 enzyme of the PPC	PPC (MP)	DNA/EST
	56710 ⁴	ptDUP	<i>P. tricornutum</i> de-ubiquitinating enzyme of the PPC	PPC	DNA/EST
	21083 ²	hCdc48	Putative cell cycle control protein Cdc48	C	DNA/EST
	20914 ²	sCdc48-1	SELMA, AAA-ATPase	PPC	DNA/EST
	19162	sCdc48-2	SELMA, AAA-ATPase	PPC^{P,F}	DNA/EST^C
	36670	hPub	Thioredoxin and PUB domain containing protein	C	DNA/EST
	37661	sPub	Thioredoxin and PUB domain containing protein	PPC^{P,F}	DNA/EST^C
	35868	alpha1	N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST
	16313	alpha2	N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST
	5613	alpha3	N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST
	9499	alpha4	N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST
	13508	alpha5	N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA
	31520	alpha6	N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST
	11025	alpha7	N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST
	38228	beta1	Catalytically active N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST
	19341	beta2	Catalytically active N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST
	15631	beta3	N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST
	866	beta4	N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST
	35268	beta5	Catalytically active N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST
	19116	beta6	N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST
	35532	beta7	N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA

	22517	beta5	N-terminal nucleophile (Ntn)-hydrolase superfamily	C	DNA/EST*
	bd_206	salpha2	N-terminal nucleophile (Ntn)-hydrolase superfamily	PPC	DNA
	45475	salpha3-1	N-terminal nucleophile (Ntn)-hydrolase superfamily	PPC^F	DNA/EST^C
	42884	salpha3-2	N-terminal nucleophile (Ntn)-hydrolase superfamily	PPC	DNA/EST
	17388	salpha7-1	N-terminal nucleophile (Ntn)-hydrolase superfamily	PPC^F	DNA/EST^D
	43079	salpha7-2	N-terminal nucleophile (Ntn)-hydrolase superfamily	PPC^F	DNA/EST*^D
	35028	sbeta2	N-terminal nucleophile (Ntn)-hydrolase superfamily	PPC^F	DNA/EST^D
	12731	sbeta3	N-terminal nucleophile (Ntn)-hydrolase superfamily	PPC	DNA
	49432	sbeta6	N-terminal nucleophile (Ntn)-hydrolase superfamily	PPC^P	DNA/EST^D
	22110	sbeta7	N-terminal nucleophile (Ntn)-hydrolase superfamily	PPC^F	DNA/EST^D
	28219	Rpt1	ATP-dependent subunit of the base subcomplex	C	DNA/EST
	18468	Rpt2	ATP-dependent subunit of the base subcomplex	C	DNA/EST
	19787	Rpt3	ATP-dependent subunit of the base subcomplex	C	DNA/EST
	45122	Rpt4	ATP-dependent subunit of the base subcomplex	C	DNA/EST
	42453	Rpt5	ATP-dependent subunit of the base subcomplex	C	DNA/EST
	30067	Rpt6	ATP-dependent subunit of the base subcomplex	C	DNA/EST
	13622	Rpn1	26S proteasome regulatory complex component	C	DNA/EST
	19300	Rpn2	26S proteasome regulatory complex component	C	DNA/EST
	10321	Rpn10	26S proteasome regulatory complex component	C	DNA/EST
	52825	Rpn3	26S proteasome regulatory complex component	C	DNA/EST
	24865	Rpn5	26S proteasome regulatory complex component	C	DNA/EST
	54715	Rpn6	26S proteasome regulatory complex component	C	DNA*
	32994	Rpn7	26S proteasome regulatory complex component, structural integrity	C	DNA/EST
	35065	Rpn9	26S proteasome regulatory complex component	C	DNA/EST
	20434	Rpn12	26S proteasome regulatory complex component	C	DNA*
	20227	Rpn11	26S proteasome regulatory complex component, substrate deubiquitination	C	DNA/EST
	17259	Rpn8	26S proteasome regulatory complex component	C	DNA*
Proteasen	37032	ClpS	ATP-dependent Clp protease adaptor protein	C	DNA
	33226	ClpS	ATP-dependent Clp protease adaptor protein	P	DNA/EST
	32782	ClpS	ATP-dependent Clp protease adaptor protein	P	DNA/EST

Chaperone	37254	Tlp1	Trypsin-like serine protease	C	DNA/EST*
	23414	sTlp1	Trypsin-like serine protease	PPC^P	DNA/EST^D
	35803	sTlp2	Trypsin-like serine protease	PPC^F (MP)	DNA/EST^C
	40345	sTlp3	Trypsin-like serine protease	PPC^F (MP)	DNA/EST^D
	40553	sTlp4	Trypsin-like serine protease	PPC^F	DNA^C
	44903	sTlp5	Trypsin-like serine protease	PPC/ER	DNA/EST
	49698	sTlp6	Trypsin-like serine protease	PPC^F	DNA/EST^D
	23486	FtsH	ATP-dependent Zn protease FtsH	M	DNA/EST
	40811	FtsH	ATP-dependent Zn protease FtsH	M	DNA/EST
	33084	FtsH	ATP-dependent Zn protease FtsH	P (MP)	DNA/EST
	9574	FtsH	ATP-dependent Zn protease FtsH	P	DNA/EST*
	19855	TCP1	Chaperonin-like superfamily (alpha SU)	C	DNA/EST
	39756	TCP1	Chaperonin-like superfamily (beta SU)	C	DNA/EST
	37901	TCP1	Chaperonin-like superfamily (gamma SU)	C/M	DNA/EST
	16096	TCP1	Chaperonin-like superfamily (delta SU)	C	DNA/EST
	42108	TCP1	Chaperonin-like superfamily (epsilon SU)	C	DNA/EST
	15395	TCP1	Chaperonin-like superfamily (zeta SU)	C	DNA/EST
	51027	TCP1	Chaperonin-like superfamily (eta SU)	C	DNA/EST
	23741	TCP1	Chaperonin-like superfamily (theta SU)	C	DNA/EST
	18019	GroEL	Cpn60-like	P	DNA/EST
	41695	GroEL	Cpn60-like	M	DNA/EST
	46754	DnaJ	DnaJ Heat shock protein	P	DNA/EST
	18036	DnaJ	DnaJ Heat shock protein	ER/M	DNA/EST
	32819	DnaJ	DnaJ Heat shock protein (Hsp40A)	C	DNA/EST
	20053	DnaJ	DnaJ Heat shock protein	C	DNA/EST
	32593	DnaJ	DnaJ Heat shock protein	ER/M	DNA/EST
	39158	DnaJ	DnaJ Heat shock protein	C (MP)	DNA/EST
	48633	sDpc	DnaJ and PDI domain containing protein	PPC^P	DNA/EST^D
	34512	sDtc	DnaJ and TPR domain containing protein	PPC^F	DNA/EST^D
	14328	DnaJ	DnaJ Heat shock protein	C	DNA/EST*

21783	DnaJ	DnaJ Heat shock protein	C	DNA/EST
45806	DnaJ	DnaJ Heat shock protein	M	DNA/EST*
39823	DnaJ	DnaJ Heat shock protein	C	DNA/EST
36016	DnaJ	DnaJ and ZnF_C2H2 domain	C	DNA/EST
33904	DnaJ	DnaJ Heat shock protein	C	DNA/EST
33781	DnaJ	DnaJ Heat shock protein	C	DNA
32201	DnaJ	DnaJ Heat shock protein	C (MP)	DNA/EST*
35283	DnaJ	DnaJ Heat shock protein; transmembrane Tmemb_185a family	ER (MP)	DNA
39170	DnaJ	DnaJ Heat shock protein	C	DNA/EST
53951	DnaJ	DnaJ Heat shock protein	C	DNA
36045	DnaJ	DnaJ Heat shock protein	ER	DNA/EST
31605	DnaJ	DnaJ Heat shock protein	C	DNA
35816	DnaJ	DnaJ Heat shock protein	C	DNA
22848	DnaJ	DnaJ Heat shock protein	C	DNA/EST*
41366	DnaJ	DnaJ Heat shock protein	C	DNA/EST
55122	-	Heat shock protein, similar to Hsp70	ER	DNA/EST
54019 ²	hHsp70	Heat shock protein 70	C	DNA/EST
55890 ¹	sHsp70	Heat shock protein 70	PPC	DNA/EST
54246	BiP	Heat shock protein 70	ER	DNA/EST
50480	sDnaK	Heat shock protein 70, prokaryotic type	PPC^F	DNA/EST^D
41172	Calr	Calreticulin, calnexin/calreticulin cycle-ER quality control	ER	DNA/EST
42287	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	C	DNA/EST
10168	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	C	DNA/EST
18458	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	ER	DNA/EST
19751	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	P	DNA/EST
19107	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	ER/P	DNA*
41671	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	ER/M	DNA*
10571	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	M	DNA*
9070	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	C	DNA
34379	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	C	DNA/EST

Vesikel-Faktoren	38525	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	C	DNA
	33225	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	C/ER	DNA/EST
	41733	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	C	DNA/EST
	35660	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin), Ubox (modified RING dom.)	C/M	DNA/EST
	16025	-	Peptidyl-prolyl cis-trans isomerase (cyclophilin)	C	DNA/EST
	32252	PDI	Protein disulfide isomerase-related	C (MP)	DNA/EST
	38945	PDI	Protein disulfide isomerase-related	C (MP)	DNA
	31858	PDI	Protein disulfide isomerase	ER (MP)	DNA/EST
	34486	PDI	Protein disulfide isomerase-related	ER	DNA/EST
	47306	PDI	Protein disulfide isomerase	ER	DNA/EST
	31585	PDI	Protein disulfide isomerase	ER/M	DNA/EST*
	35302	Ero1	Thiol oxidase required for oxidative protein folding in the ER	ER	DNA/EST
	22713	Rab1a	ER to Golgi transport	C	DNA/EST
	41867	Rab1b	ER to Golgi transport	ER/C^P	DNA*^D
	14604	Rab2	Cis-Golgi and ER to Golgi transport	C	DNA/EST
	18141	Rab5a	Endocytosis/endocytic-sorting	C	DNA/EST
	44567	Rab5b	Endocytosis/endocytic-sorting	C	DNA/EST*
	53175	Rab5c	Endocytosis/endocytic-sorting	C	DNA
	21535	Rab6	Endosomes to Golgi as well as retrograde transport to the ER	C	DNA/EST
	51199	Rab7a	Transport from early to late endosomal stages	C	DNA/EST
	40171	Rab7b	Transport from early to late endosomal stages	C	DNA
	30139	Sec4	Post-Golgi transport to plasma membrane	C	DNA/EST
	31160	Rab11	Recycling from endosome to trans-Golgi and plasma membrane	C	DNA/EST
	44216	Rab11/RabX2	Recycling from endosome to trans-Golgi and plasma membrane	C	DNA/EST
	51810	Rab18	Endocytic transport	C	DNA
	37370	Bet2	Geranylgeranyltransferase type II (GGTase-II)-like protein	C	DNA/EST
	24080	-	Protein farnesyltransferase (FTase)-like protein	C	DNA/EST
	35651	Bet4	Protein farnesyltransferase (FTase)-like protein	C	DNA/EST
	46321	Pra1	Rab guanine dissociation inhibitor (GDI) displacement factor	ER (MP)	DNA/EST*
	47212	Pra1	Rab guanine dissociation inhibitor (GDI) displacement factor	C (MP)	DNA*

	22307	Gdi1	GDI (GDP dissociation inhibitor)	C	DNA/EST
	43871	Gdi2	GDI (GDP dissociation inhibitor)	M	DNA*
	19985	Gyp1	RabGAP/TBC	C/M	DNA/EST
	39847	-	RabGAP/TBC	C	DNA/EST
	35506	YPL196Wp-like	RabGAP/TBC	C	DNA/EST
	10479	Gyp5	RabGAP/TBC	C	DNA*
	55114	-	RabGAP/TBC	C	DNA*
	38859	-	RabGAP/TBC	C	DNA/EST
	38862	-	RabGAP/TBC	C	DNA/EST
	46730	Rap/Ran-GAP	Rap/Ran-GAP	C/M	DNA
	56064	Sec7-like Arf-GEF	Sec7-like Arf-GEF	C	DNA
	43545	-	Sec7-like Arf-GEF	C	DNA*
	54143	-	Sec7-like Arf-GEF	C	DNA*
	50615	Sec7C	Sec7-like Arf-GEF	C	DNA
	54591	Sec7D	Sec7-like Arf-GEF	C	DNA
	38808	Yip1	Rab GTPase interacting factor, Golgi membrane protein	C (MP)	DNA/EST
	38327	Yip1	Rab GTPase interacting factor, Golgi membrane protein	C (MP)	DNA*
COPI	43251	ARF1	Arf-type small GTPase, probable Arf1	C	DNA/EST
	21929	COPalpha	COPI coatomer protein complex, subunit alpha	C/M	DNA/EST
	54511	COPbeta	Beta coatomer, subunit of the coat protein complex COP I	C	DNA/EST
	27518	COPbeta2	Beta prime coatomer, subunit of the coat protein complex COP I	C	DNA/EST
	30436	COPdelta	Delta coatomer, subunit of the coat protein complex COP I	C/M	DNA/EST
	38981	COPzeta	Zeta coatomer, subunit of the coat protein complex COP I	C	DNA
	18537	COPgamma	Gamma coatomer, subunit of the coat protein complex COP I	C	DNA/EST
COPII	54420	Sar1a	Sar1-like small GTPase	C/ER	DNA/EST
	15349	Sar1b	Sar1-like small GTPase	C/M	DNA
	41706	Sec13	Component of the coat protein complex COP II	C/ER	DNA/EST
	42673	Sec23	Component of the vesicle coat protein complex COP II	C	DNA/EST
	47270	Sec23-like	Component of the vesicle coat protein complex COP II	C	DNA/EST*
	36717	Sec24	Component of the vesicle coat protein complex COP II	C	DNA/EST

				C ^P	DNA/EST ^D
Clathrin-abhängige Endocytose	48220	Sec24-like	Member of the Sec24p family		
	37106	Sec31-like	Vesicle coat complex COPII, subunit Sec31-like	C	DNA/EST
	19731	APmu1	Adapter-related protein complex 1 mu subunit	C	DNA/EST
	18142	APmu2	Adapter-related protein complex 2 mu subunit	C	DNA/EST
	55392	APmu3	Adapter-related protein complex 3 mu subunit	C	DNA/EST
	45284	APmu4	Adapter-related protein complex 4 mu subunit	C	DNA/EST*
	36149	AP1sigma	Adapter-related protein complex 1 sigma subunit	C	DNA/EST
	54935	AP2sigma	Adapter-related protein complex 2 sigma subunit	C	DNA/EST
	14658	AP3sigma	Adapter-related protein complex 3 sigma subunit	C	DNA
	54718	AP4sigma	Adapter-related protein complex 4 sigma subunit	C	DNA/EST
	46599	CLC	Clathrin light chain	C	DNA/EST
	54801	CHC	Clathrin heavy chain	C	DNA/EST
	13511	AP1gamma	Adapter-related protein complex 1 subunit gamma	C	DNA/EST
	1735	AP1/2beta	Adapter-related protein complex 1/2 subunit beta	C	DNA/EST
	54442	AP2alpha	Adapter-related protein complex 3 subunit alpha	C	DNA/EST
	54028	AP3delta	Adapter-related protein complex 3 subunit alpha	C	DNA/EST
	54789	AP3beta	Adapter-related protein complex 3 subunit beta	C	DNA/EST*
	54730	AP4beta	Adapter-related protein complex 4 subunit beta	C	DNA/EST
	41331	AP4epsilon	Adapter-related protein complex 4 subunit epsilon	C	DNA/EST
SNAREs	36912	SytB	Putative syntaxin, t-SNARE component, similar to plant Syp41	C/M (MP)	DNA/EST
	43244	-	Putative vesicle transport involved protein, similar to plant Vti12	C (MP)	DNA/EST
	40521	-	v-SNARE, similarity to plant vesicle transport protein Vti12	C (MP)	DNA/EST
	48591	SybC	Putative synaptobrevin, v-SNARE component	C (MP)	DNA/EST
	54855	-	Putative SNARE, similarity to yeast v-SNARE and plant t-SNAREs	C (MP)	DNA/EST
	43259	-	Syntaxin 18-like	C (MP)	DNA/EST
	17683	SybA	Putative synaptobrevin, v-SNARE, similar to plant Vamp711/13/14	ER^F (MP)	DNA/EST^C
	31559	-	Syntaxin, t-SNARE; similar to plant Syp61	C (MP)	DNA
	55241	-	Similar to syntaxin 5	C (MP)	DNA/EST
	54066	SybE	Putative synaptobrevin, v-SNARE, similar to Sec22	C (MP)	DNA/EST

ESCRT	43261	-	t-SNARE component	C	DNA/EST*
	44364	GOS	Putative Golgi SNARE, v-SNARE protein with similarity to AtGOS	C (MP)	DNA/EST
	44266	-	SNARE associated protein	C (MP)	DNA/EST
	44804	-	t-SNARE component, similar to plant Syp71	C (MP)	DNA/EST
	45601	SybD	Putative synaptobrevin, v-SNARE, similar to plant Vamp722	ER/C^F (MP)	DNA/EST^C
	37104	Bet1	t-SNARE component, similar to Bet1	C (MP)	DNA/EST
	51073	SybF	Putative synaptobrevin, similar to plant Ykt61	C	DNA/EST
	49105	-	Similar to vesicle transport v-SNARE protein, Vti12	C	DNA/EST
	52139	Memb11	Putative v-SNARE, similar to plant Membrin11 and animal GOSR	C (MP)	DNA
	45881	Vam3	t-SNARE component, similar to Vam3	C/M	DNA/EST
	41159	-	Putative synaptobrevin, similar to plant Vamp714	C (MP)	DNA/EST
	31234	-	SNARE associated protein	P/M (MP)	DNA
	41130	-	SNARE associated protein	C (MP)	DNA/EST
	43839	-	SNARE associated protein	C (MP)	DNA/EST
	44014	-	SNARE associated protein	ER/P (MP)	DNA/EST
	45676	-	SNARE associated protein	C (MP)	DNA
	48103	SNAP	Putative soluble NSF attachment protein ortholog (alpha-SNAP)	C	DNA/EST
	41073	NSF	N-ethylmaleimide-sensitive factor, similar to Sec18	C	DNA/EST
	36682	NSF	N-ethylmaleimide-sensitive factor, similar to Sec18	P/M	DNA
	23807	NSF	N-ethylmaleimide sensitive factor	P	DNA/EST
	31836	NSF	N-ethylmaleimide sensitive factor	P	DNA/EST
	20244	NSF	N-ethylmaleimide sensitive factor	P	DNA/EST
	31375	SNF7/DID2	Endosomal sorting complex required for transport III, SNF7-family	C	DNA/EST
	5904	SNF7/DID4	Endosomal sorting complex required for transport III, SNF7-family	C	DNA/EST
	12752	Doa4-like	Peptidase C19, ubiquitinyl hydrolase	C	DNA*
	3692	Doa4-like	Peptidase C19, ubiquitinyl hydrolase	C	DNA/EST*
	974	Doa4-like	Peptidase C19, ubiquitinyl hydrolase	C	DNA/EST*
	38130	SNF7-like	Similar to vacuolar sorting protein SNF7, SNF7-family	C	DNA/EST
	18620	Vps60-like/SNF-like	Similar to vacuolar sorting protein SNF7	C	DNA/EST
	48646	SNF7-like	Similar to vacuolar sorting protein SNF7	C	DNA/EST

Cytoskelett	23040	DID4-like	SNF7-family	C	DNA/EST
	30898	Vps4	AAA-ATPase involved in multivesicular body protein sorting	C	DNA/EST
	49220	Vps20	Myristoylated subunit of ESCRT-III	C	DNA/EST
	16493	Vps22/SNF8	Component of the ESCRT-II complex	C/M	DNA/EST*
	28874	Vps23	Ubiquitin receptor Vps23	C	DNA/EST
	38893	Vps25	Similar to ESCRT-II complex component Vps25	C	DNA
	34801	Vps28	Component of the ESCRT-I complex	C	DNA/EST*
	16608	Vps27	Endosomal protein that forms a complex with Hse1p	C	DNA/EST*
	51157	Act1	Actin/Actin-like protein	C	DNA/EST
	42173	Act2	Actin/Actin-like protein	C	DNA/EST
	10514	ARP1	Actin related protein	C	DNA/EST
	34772	-	Actin related protein	C	DNA/EST
	36432	ARP	Actin related protein	C	DNA/EST
	32362	-	Actin related protein	C	DNA
	20557	TUB1	Alpha tubulin	C	DNA/EST
	13540	TUB2	Beta Tubulin	C	DNA/EST
	10652	TUB3	Gamma Tubulin	C	DNA/EST
	37279	Myo29	Myosin 29	C	DNA
	ACS35537.1	MyoA	Myosin A	C	DNA/EST
	36456	MyoB	Myosin B	C	DNA
	ACS35539.1	MyoC	Myosin C	C	DNA
	34387	MyoD	Myosin D	C	DNA
	ACS35541.1	MyoE	Myosin E	C	DNA/EST
	33529	MyoF	Myosin F	C	DNA
	19874	MyoG	Myosin G	C	DNA/EST
	40705	MyoH	Myosin H	C	DNA
	37190	MyoI	Myosin I	C	DNA
	49425	Ank1	Ankyrin repeat family protein	P	DNA/EST
	40127	Ank2	Ankyrin repeat family protein	ER/PPC	DNA/EST
	36814	Ank3-1	Ankyrin repeat family protein	ER/PPC	DNA/EST

	50055	Ank3-2	Ankyrin repeat family protein	ER/PPC	DNA/EST
	37621	Ank4	Ankyrin repeat family protein	ER/PPC	DNA/EST
	41153	Ank5	Ankyrin repeat family protein	M^F	DNA/EST^C
	31896	Ank6	Ankyrin repeat family protein	C	DNA/EST
	22063	-	Ankyrin repeat family protein	C	DNA*
	35186	-	Ankyrin repeat family protein	C	DNA/EST
	38557	-	Ankyrin repeat family protein	C	DNA/EST
	36555	-	Ankyrin repeat family protein	M	DNA*
	38169	-	Ankyrin repeat family protein	C/M	DNA/EST
	35740	-	Ankyrin repeat family protein, DnaJ domain	C	DNA/EST
	31883	-	Ankyrin repeat family protein	C	DNA/EST*
Plastidenteilungs- apparat	55912	Dyn1	Dynamin participates in the endocytosis	C	DNA/EST
	37379	sDrp (DRP5B)	Dynamin-related GTPase	PPC^{P,F}	DNA*^C
	38279	FtsZ1	Procaryotic cell division GTPase	P	DNA/EST*
	51264	FtsZ2	Procaryotic cell division GTPase	P	DNA/EST
	38810	FtsZ3	Procaryotic cell division GTPase	P	DNA/EST
Glycogenine	40510	GT8-1	Glycogenin-like protein	ER/PPC	DNA*
	40314	sPdr1 (GT8-2)	Glycogenin-like protein	PPC^F	DNA^C
	39832	GT8-3	Glycogenin-like protein	C	DNA
	38120	GT8-4	Glycogenin-like protein	ER	DNA*
	35314	GT8-5	Glycogenin-like protein	ER	DNA*
Zwei-Komponenten System	50334	SHK	Probable two-component sensor histidine kinase	P/M	DNA/EST
	27129	DRR1	Two-component response regulator; Diatom Response Regulator 1	P/M	DNA/EST
	35261	DRR2	Two-component response regulator; Diatom Response Regulator 2	P	DNA*
	38264	DRR3	Two-component response regulator; Diatom Response Regulator 3	C	DNA
Nucleosid- diphosphatkinasen	52812	NDK1	Nucleoside Diphosphate Kinase 1	P	DNA
	42104	NDK2	Nucleoside Diphosphate Kinase 2	C	DNA
	18001	NDK3	Nucleoside Diphosphate Kinase 3	M^F	DNA/EST^C

Kinasen	35463	Aurora-A	Aurora-A like protein kinase	C	DNA
	8773	MARK3	MAP/microtubule affinity-regulating kinase	C	DNA/EST
	54265	CaMK2	Calcium/calmodulin-dependent protein kinase	C	DNA*
	34071	Chk2	Chk2-loki isoform A, homolog of the yeast Mek1p	C	DNA/EST
	12820	CDPK1	Calcium dependent protein kinase of the CaMK subfamily	C	DNA/EST*
	20262	CDKA1/CAK2	Homolog CDK7, CTD kinase activity	C	DNA/EST
	10160	CDKD1/CAK4	CAK4, CTD kinase activity	C	DNA*
	51279	CDKA2	Putative B-type CDK involved in G2-M transition	C	DNA/EST
	34165	CDK3	Hypothetical cyclin dependent kinase 3	C	DNA*
	49025	CDK4	Hypothetical cyclin dependent kinase 4	C	DNA
	23198	CDK5	Cyclin dependent kinase 5	C	DNA/EST
	5526	CDK6	Hypothetical cyclin dependent kinase 6	C	DNA/EST*
	14453	CDK7	Hypothetical cyclin-dependent kinase 7	C	DNA/EST*
	54091	CK2_1	Protein CK2 catalytic subunit alpha	C	DNA/EST
	44716	MYT1	MYT1 inhibitory kinase	C	DNA
	17550	STK1	Serine/Threonine Kinase-like	ER/P	DNA/EST
	34972	STK2	Serine/Threonine Kinase-like	C	DNA/EST
	32111	STK3	Serine/Threonine Kinase-like	C	DNA*
	33437	STK4	Serine/Threonine Kinase-like	ER/PM^P	DNA/EST^D
	10628	Polo-like	Polo-like kinases	C	DNA*
	40278	Pk	Protein kinase	ER^F	DNA^C
Cryptochrome	27429	CPF1	Cryptochrome/photolyase family	C	DNA/EST
	34592	CPF2	Cryptochrome/photolyase family	ER^P	DNA^D
	54342	-	Cryptochrome/photolyase family-like	P/M	DNA/EST*
	36159	CPD1	Cyclobutane pyrimidine dimer 1, deoxyribodipyrimidine photolyase family	C	DNA/EST
	55091	-	Cryptochrome/photolyase family-like	C/M	DNA
Glutathion-System	50476	GltP	Glutathione peroxidase	ER^P	DNA/EST^D
	36273	GPX	Glutathione peroxidase	ER	DNA*
	39148	-	Glutathione peroxidase domain-containing protein	C/M	DNA
	52119	-	Glutathione peroxidase domain-containing protein	ER (MP)	DNA/EST

	32717	GRX	Glutaredoxin	C	DNA*
	18564	-	Glutaredoxin	M (MP)	DNA*
	35033	-	Glutaredoxin	P	DNA
	49167	GLR/Gsr1	Glutathione reductase	C	DNA/EST
	20349	GLR/Gsr2	Glutathione reductase	M^P	DNA/EST^D
	21542	-	Glutathione reductase, Thioredoxin reductase	C/M	DNA/EST
Lipidtransfer	35683	Sec14-1	Sec14p-like protein, PtdIns/PtdCho transfer protein	C	DNA/EST
	31506	Sec14-2	Sec14p-like protein, PtdIns/PtdCho transfer protein	C	DNA/EST
	36050	Sec14-3	Sec14p-like protein, PtdIns/PtdCho transfer protein	C	DNA/EST
	29506	Sec14-4	Sec14p-like protein, PtdIns/PtdCho transfer protein	C	DNA/EST
	44766	sSec14	Sec14p-like protein	PPC^F	DNA/EST^C
Fettsäure- Biosynthese	36633	FabH	3-oxoacyl-[acyl-carrier-protein] synthase III	U	DNA
	37367	FabFa	3-oxoacyl-[acyl-carrier-protein] synthase II	P	DNA
	18940	FabFb	3-oxoacyl-[acyl-carrier-protein] synthase II	M	DNA/EST
	42339	FabB	3-oxoacyl-[acyl-carrier-protein] synthase I	P	DNA/EST
	20902	FabG	3-oxoacyl-[acyl-carrier protein] reductase	P	DNA/EST
	33354	-	Short chain dehydrogenase	C	DNA/EST
	10567	-	3-oxoacyl-[acyl-carrier protein] reductase	C	DNA*
	38987	FabZ1	DnaJ domain-containing protein	C	DNA/EST
	55157	FabZ2	3R-hydroxymyristoyl ACP dehydrase	P	DNA/EST
	18109	FabI	Enoyl-[acyl-carrier protein] reductase I	P	DNA/EST
	21723		Enoyl-[acyl-carrier protein] reductase I	M	DNA/EST*
	21626	FabD	[acyl-carrier-protein] S-malonyltransferase	P	DNA/EST
	32224	-	Stearoyl-ACP desaturase	P	DNA/EST
	46830	PTD5a	Delta5 fatty acid desaturase	C (MP)	DNA/EST
	15012	PTD5b	Delta5 fatty-acid-desaturase	C (MP)	DNA/EST
	25769	FAD2	Delta12-fatty acid dehydrogenase	C (MP)	DNA/EST
	29488	PTD6	Delta6-fatty acyl-CoA desaturase	C (MP)	DNA/EST
	22677	-	Dihydroceramide delta-4 desaturase	C	DNA/EST

Glycerophospholipid- /Glycerolipid- Metabolismus	28797	PTD9	Stearoyl-CoA 9-desaturase	ER (MP)	DNA/EST
	55137	-	Delta6-fatty acyl-CoA desaturase	C (MP)	DNA/EST
	50443	-	Delta6-fatty acyl-CoA desaturase	P/ER	DNA/EST
	46383	-	Delta4-sphingolipid fatty acid desaturase	P	DNA/EST
	38997	-	Delta6-fatty acyl-CoA desaturase	ER (MP)	DNA/EST
	44622	-	Delta6-fatty acyl-CoA desaturase	P	DNA/EST
	48423	PTD12	Omega-6 fatty acid desaturase	P	DNA/EST
	32140	-	CDP-diacylglycerol-inositol 3-phosphatidyltransferase	ER	DNA/EST
	28209	GPD1	Glycerol-3-phosphate dehydrogenase (NAD+)	U	DNA/EST*
	32033		Glycerol-3-phosphate dehydrogenase (NAD+)	M	DNA/EST*
	38538	GlpA	Glycerol-3-phosphate dehydrogenase	C	DNA*
	54709	GPT1	Glycerol-3-phosphate o-acyltransferase	ER (MP)	DNA/EST
	37991	DGK	Diacylglycerol kinase	C (MP)	DNA/EST
	39977	CDS1	Phosphatidate cytidyltransferase	M	DNA/EST
	37411		Phosphatidate cytidyltransferase	ER (MP)	DNA/EST*
	37772	CDS1	Phosphatidate cytidyltransferase	ER (MP)	DNA/EST
	47576	PgsA	CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase	M	DNA/EST*
	31694		CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase	ER (MP)	DNA
	33864	EPT1/CPT1	Ethanolaminephosphotransferase/diacylglycerol cholinephosphotransferase	M	DNA
	37086		Ethanolaminephosphotransferase/diacylglycerol cholinephosphotransferase	C (MP)	DNA*
	12431	PLD	Phospholipase D	C	DNA/EST*
	40261	PPAP2	Phosphatidate phosphatase	M	DNA/EST
	39949		Phosphatidate phosphatase	C (MP)	DNA
	42871	PEMT	Phosphatidylethanolamine N-methyltransferase	C (MP)	DNA/EST
	20460	LPT1	Lysophospholipid acyltransferase	C (MP)	DNA/EST
	32364	LCAT	Lecithin-cholesterol acyltransferase	M	DNA/EST
	49702		Lecithin-cholesterol acyltransferase	M	DNA/EST
	35781	PLA2G	Phospholipase A2	C	DNA

	21816	Lypla1	Lysophospholipase I	M	DNA/EST
	9280	ADPRM	Manganese-dependent ADP-ribose/CDP-alcohol diphosphatase	M	DNA/EST
	23235	PCYT2	Ethanolamine-phosphate cytidyltransferase	C	DNA/EST
	44584	PCYT1	Choline-phosphate cytidyltransferase	M	DNA/EST
	48664	EKI1	Ethanolamine kinase	C	DNA/EST
	46453	CHK	Choline kinase	M	DNA/EST
	48153	GlxK	Glycerate kinase	M	DNA/EST
	18674	GlpK	Glycerol kinase	C	DNA/EST*
	49544	MGAT	Acyl-CoA:2-acylglycerol O-acyltransferase	ER (MP)	DNA/EST
	43469		Acyl-CoA:2-acylglycerol O-acyltransferase	ER	DNA/EST
	31662		Acyl-CoA:2-acylglycerol O-acyltransferase	C (MP)	DNA
	49462		Acyl-CoA:2-acylglycerol O-acyltransferase	M	DNA/EST
	32834	PNPLA3	Triacylglycerol lipase	C	DNA
	50397		Triacylglycerol lipase	M	DNA/EST
	47488		Triacylglycerol lipase	ER	DNA/EST
Carboanhydrasen	44526	α Ca1	Alpha carbonic anhydrase	ER ^P	DNA/EST ^D
	54251	α Ca2	Alpha carbonic anhydrase	ER ^P	DNA/EST ^D
	55029	α Ca3	Alpha carbonic anhydrase	ER/P	DNA/EST
	35370	α Ca1	Alpha carbonic anhydrase	PPC ^P	DNA ^D
	42574	α Ca2	Alpha carbonic anhydrase	PPC ^P	DNA/EST ^C
	51305	β Ca1 (PtCA1)	Beta carbonic anhydrase	P	DNA/EST
	45443	β Ca2 (PtCA2)	Beta carbonic anhydrase	P	DNA/EST
	36906	γ Ca1	Gamma carbonic anhydrase-like	C/M	DNA
	50900	γ Ca2	Gamma carbonic anhydrase-like	C/M	DNA/EST
	41375	γ Ca3	Gamma carbonic anhydrase-like	C/M	DNA/EST
<i>G. theta</i> Nm Homologe	31704	sORF139	Similar to <i>Guillardia theta</i> nucleomorph encoded ORF139	PPC ^F	DNA ^C
	47811	sORF261	Similar to <i>G. theta</i> nucleomorph encoded ORF261 / <i>Cyanidioschyzon merolae</i> unknown GTP-binding protein CMD101C (RHD3/Sey1 family-like GTPase)	PPC ^F	DNA/EST* ^D
	45935	sORF532a	Similar to <i>G. theta</i> nucleomorph encoded ORF532a / <i>C. merolae</i> CMQ054C (RHD3/Sey1 family-like GTPase)	PPC ^F	DNA/EST ^{C,D}

	48879	sORF534	Similar to <i>G. theta</i> nucleomorph encoded ORF534/<i>C. merolae</i> CMI197C (RHD3/Sey1 family-like GTPase)/ <i>T. gondii</i> PPP1 (apicoplast protein import)	PPC^F	DNA^{C,D}
	49840	ORF387	Similar to <i>G. theta</i> nucleomorph encoded ORF387/<i>C. merolae</i> CMP057C	ER/N^F	DNA/EST^D
Prolyl 4-Hydroxylasen	44433	P4H-1	Prolyl 4-hydroxylase domain containing protein	C	DNA/EST
	44303	P4H-2	Prolyl 4-hydroxylase domain containing protein	C/M	DNA
	47766	sP4h	Prolyl 4-hydroxylase and TPR domain containing protein	PPC^P	DNA^C
Pectinestarasen	54713	-	Pectin esterase-like domain containing protein, PT repeats	ER	DNA*
	47888	-	Pectin esterase-like domain containing protein, PDZ domain, PT repeats	C (MP)	DNA
	48565	-	Pectin esterase-like domain containing protein, PT repeats	C (MP)	DNA*
	49133	-	Pectin esterase-like domain containing protein, PT repeats	ER	DNA
	47701	-	Pectin esterase-like domain containing protein, PT repeats	C (MP)	DNA
	54178	-	Glycoprotein precursor, Pectin esterase-like domain containing protein	C (MP)	DNA
	47315	-	Pectin esterase-like domain containing protein, PT-repeats	ER/P	DNA/EST
	46569	-	Pectin esterase-like domain containing protein, PT repeats	ER/M (MP)	DNA*
	47646	-	Pectin esterase-like domain containing protein, PT repeats	C (MP)	DNA/EST
	45068	-	Pectin esterase-like domain containing protein, PT repeats	ER	DNA/EST
	42459	-	Pectin esterase-like domain containing protein, PT repeats	ER	DNA
	50564	-	Pectin esterase-like domain containing protein, PT repeats	ER	DNA/EST
	49933	-	Pectin esterase-like domain containing protein, PT repeats	ER	DNA/EST
	37424	sPel	Pectin esterase-like domain containing protein, PT-repeats	PPC^P	DNA^C
Frustuline	38415	Fru	Frustulin-like	ER^P	DNA/EST^D
	48054	Fru2	Frustulin, cell wall assoicated protein	ER/P	DNA/EST
	51797	Fru3	Frustulin, cell wall assoicated protein	ER	DNA/EST*
	48049	TRD6	Frustulin-like,contains a tryptophan rich domain (TRD)	ER/P	DNA/EST
	54863	TRD1	Frustulin-like,contains a tryptophan rich domain (TRD)	ER	DNA/EST
	38413	TRD2	Frustulin-like,contains a tryptophan rich domain (TRD)	C	DNA*
	54558	TRD3	Frustulin-like,contains a tryptophan rich domain (TRD)	C	DNA/EST*
	49571	-	Frustulin-like, leucine-rich repeat protein	ER	DNA/EST
	47872	-	Frustulin-like,contains a tryptophan rich domain (TRD)	ER	DNA

Sonstige	38198	-	Frustulin-like,contains a tryptophan rich domain (TRD)	ER	DNA
	38417	-	Frustulin-like,contains a tryptophan rich domain (TRD)	ER	DNA
	38418	-	Frustulin-like,contains a tryptophan rich domain (TRD)	ER	DNA/EST
	51925	-	Frustulin-like	ER	DNA
	39072	-	Frustulin-like	C	DNA
	36721	G-Protein	Heterotrimeric G protein alpha subunit 1	C	DNA/EST*
	21897	G-Protein	Heterotrimeric G protein beta subunit	C	DNA/EST
	8659	ARL1	Arf subfamily small GTPase, similar to Arf-like (ARL) proteins	C	DNA
	51819	ARL2	Arf-subfamily small GTPase, similar to Arf-like proteins (ARL)	C	DNA
	13157	ARL10-like	Arf-subfamily small GTPase, similar to Arf-like proteins (ARL)	C/ER	DNA
	17300	-	Platelet-activating factor acetylhydrolase beta subunit-like protein	C	DNA/EST
	29608	PAF1	PolyAdenylation Factor subunit 1	C	DNA/EST
	45580	Coronin	Coronin like protein, actin-binding protein	C	DNA/EST
	11509	SEH1-like	Nuclear pore protein, homologous to Sec13p	C	DNA
	17948	CAF1	Chromatin assembly factor subunit c	ER/P (MP)	DNA*
	21105	MCF1	Mitochondrial carrier family protein (Nukleotide carrier)	ER/M	DNA/EST
	55928	AAA-ATPase	Nuclear valosin containing protein-like protein, Rix7-homolog	N/C	DNA/EST
	22177	14-3-3-protein	14-3-3-protein	C	DNA
	32793	-	GTP-binding protein	M	DNA/EST
	42161	-	GTP-binding protein	ER/P/M	DNA*
	31880	-	Putative E3 ubiquitin-protein ligase, SSM4-like	C (MP)	DNA
	47723	-	UDP-glucose glycoprotein glucosyltransferase	ER/M	DNA/EST
	39931	-	WD40-repeat protein	C	DNA/EST
	5170	-	Prohibitin-like protein, Band7-like domain	ER/P/M	DNA/EST
	37716	Tom70	Translocase of the outer mitochondrial membrane 70-like protein	M^P	DNA/EST^D
	37731	Ubr1	E3 ubiquitin ligase (N-recognin)-like protein	C/M	DNA*
	40578	Ubr1	E3 ubiquitin ligase (N-recognin)-like protein	C	DNA
	40547	Ubr1	E3 ubiquitin ligase (N-recognin)-like protein	C	DNA
	47503	Ubr1	E3 ubiquitin ligase (N-recognin)-like protein	C/ER	DNA*
	35153	APS	ATP sulfurylase	C	DNA/EST

32923	-	Inorganic pyrophosphatase (PPase)-like	C	DNA
38148	-	Inorganic pyrophosphatase (PPase)-like	P/M	DNA/EST
23100	-	ATP sulfurylase	P	DNA/EST
41316	sPrp	Pentapeptide repeats containing protein	PPC^P	DNA/EST^C
37961	FolC	Folypolyglutamate synthase	M^P	DNA/EST^D
48414	-	Folypolyglutamate synthase	C	DNA
26432	DLDH	Dihydrolipoyl dehydrogenase	C	DNA*
30113	DLDH	Dihydrolipoyl dehydrogenase	ER/P	DNA/EST
45488	-	Myo-inositol 2-dehydrogenase, mercuric reductase fusion protein	ER/P/M	DNA/EST
45428	-	Dihydrolipoyl dehydrogenase	P (MP)	DNA*
23384	-	Dihydrolipoyl dehydrogenase	C	DNA/EST
44172	HAP	Histidine acid phosphatase	ER^P	DNA^D
31412	-	Histidine acid phosphatase	M	DNA/EST
22127	PGP	Phosphoglycolate phosphatase	M^P	DNA/EST^D
36817	PPG1	Similar to yeast Muc1p (GPI-anchored cell surface glycoprotein)	P^P	DNA/EST^D
46711	-	Similar to yeast TIR4 (cell wall mannoprotein)	P/M	DNA/EST
42675	sSmc	SMC (structural maintenance of chromosomes)-domain containing protein	PPC^P	DNA/EST^D
19914	Sec63	Similar to endoplasmic reticulum translocon component Sec63	C/ER (MP)	DNA/EST
42877	CoaE	Dephospho-coenzyme A kinase	ER/V^F	DNA/EST^D
38588	DCP	Peptidase family M3 dipeptidyl carboxypeptidase	ER^F	DNA/EST^D
41002	E1-like	E1-like protein (ubiquitination)	P^F	DNA^C
44694	M6	M6 family Metalloprotease	ER^F	DNA/EST^C
50197	Trx	Thioredoxin-like protein	P^F	DNA/EST^D
45446	Phyco	Phycobilin oxidoreductase	P^F	DNA^C

Zeilen in Grau hinterlegt deuten Proteine an, die über den nicht gerichteten Ansatz identifiziert werden konnten. Alle restlichen Proteine wurden über Blast oder direkte Suche detektiert; fett gedruckte Zeilen kennzeichnen Proteine, deren Lokalisation über *in vivo* BTS-/FL-Lokalisationsstudien in *P. tricornutum* verifiziert wurde;

Prozess, die Proteine wurden hinsichtlich bekannter zellulärer Prozesse klassifiziert, in die sie für gewöhnlich involviert sind; ID, Proteinidentifikationsnummer (Phatr2DB oder NCBI);

Putative Funktion/Domänen, putative Proteinfunktionen wurden entweder durch bestehende Annotation, Detektion homologer Proteine, oder durch Identifizierung konservierter Domänen zugeordnet;

(Putative) Lokalisation, die Klassifizierung der Proteine hinsichtlich ihrer (putativen) Lokalisation erfolgte über *in silico* Analysen (SignalP, TargetP, BlastP), sowie über BTS-/FL-eGFP *in vivo* Lokalisationsstudien (fett gedruckte Zeilen): C, Cytosol; M, Mitochondrium; ER, Signalpeptid (ER und sekretorischer Weg); N, Nucleus; P, Plastide; PPC, periplastidäres Kompartiment; U, Unbekannt; V, Vakuole; MP, potentiell Membranprotein (TMHMM), die Detektion von Transmembrandomänen wurde bei der Lokalisationsvorhersage nicht berücksichtigt;^P, verifizierte *in vivo* Lokalisation über Präsequenz-eGFP Fusionsprotein; ^F, verifizierte *in vivo* Lokalisation über volle Länge (*full length*)-eGFP Fusionsprotein

Genmodell, die Spalte Genmodell gibt Aufschluss über die vorhandene EST-Datenlage: DNA, keine EST-Daten vorhanden, ID-Sequenz für *in silico* Vorhersage genutzt; DNA/EST, EST-Abdeckung für den 5'-Bereich des Genmodells (N-Terminus des kodierten Proteins); DNA*, keine EST-Daten für den 5'-Bereich, alternatives Startcodon (N-Terminus) wurde für die Lokalisationsvorhersage verwendet; DNA/EST*, EST-Abdeckung des 5'-Bereichs des Genmodells (N-Terminus des Proteins), alternatives, EST-belegtes Startcodon (N-Terminus) wurde für die Lokalisationsvorhersage verwendet; ^C, Gensequenz für *in vivo* Lokalisationsstudien aus cDNA amplifiziert; ^D, Gensequenz für *in vivo* Lokalisationsstudien aus genomischer DNA amplifiziert.

¹: Gould SB, Sommer MS, Kroth PG, Gile GH, Keeling PJ, Maier UG (2006) *Mol Biol Evol* 23:2413-2422.

²: Sommer MS, Gould SB, Lehmann P, Gruber A, Przyborski JM, Maier UG (2007) *Mol Biol Evol* 24:918-928.

³: Hempel F, Bullmann L, Lau J, Zauner S, Maier UG (2009) *Mol Biol Evol* 26:1781-1790.

⁴: Hempel F, Felsner G, Maier UG (2010) *Mol Microbiol* 76:793-801.

⁵: Bullmann L, Haarmann R, Mirus O, Bredemeier R, Hempel F, Maier UG, Schleiff E (2010) *J Biol Chem* 285(9):6848-56

Tabelle modifiziert nach Moog et al. (2011)

8.2. BTS-Sequenzen PPC-lokalisierter Faktoren aus *P. tricornutum*

Die Länge der TPL-Sequenzen wurde anhand konservierter Bereiche eines jeweiligen Proteins kalkuliert und entspricht hier dem minimalen, nicht konservierten Bereich zwischen SP (Vorhersage durch SignalP 3.0) und ermitteltem maturen Protein (NCBI BlastP). In *in vivo* Lokalisationsstudien wurden zum Teil erweiterte (längere) bzw. volle Länge Proteinsequenzen als eGFP-Fusionskonstrukte exprimiert, um in unklaren Fällen die Vollständigkeit und Integrität der N-terminalen Zielsteuerungsinformationen zu gewährleisten.

Tab. 8.2 BTS-Sequenzen PPC-lokalisierter Faktoren aus *P. tricornutum* ermittelt über Signalpeptidvorhersage (SignalP) und Detektion N-terminaler Extensionen (BlastP).

Protein	Putative BTS
sCdc48-2	<u>MRFSEAAAALTATVLLAVSVRENHGS</u> VGHAHQASSVTSTRTSKRTRVPFWGSRLDHIYTADATVTTSAILSTRGGGAPGEVRKEE...
sPub	<u>MHGLGVACALWLASAVDA</u> GLHQSPFLFRPVLQYIRPRLTSAILSVSRGGNQEEGTLEK...
sα2	<u>MLHISNCFTLVLLFGIRVSNHCAA</u> SDSADGRHSFS...
sα3-1	<u>MKTRSMVLWKAASALLLASVLTSSGSS</u> APWQDTRTPGTGWQ...
sα3-2	<u>MRVETATLQMESRTRVHRIRRSNPLTVSLILLVLLCGASPSRGA</u> AATDDTRLPLHGVSPDRFVQ...
sα7-1	<u>MHARHHMKSSGKRRHRFPDRNRVHDPMYQVIFLCLCLAKTWTLVHA</u> AHSRQRVVVRTVA...
sα7-2 (SA)	MARFRTQKRDRFQMGARGENGEYSVRGRYSQSGPMGFSLGLTIILAWIALGWPSVAAKSASNISPIRNFEPSLAASSAQGTTISVSY RDGSRQQTG/VVV...
sβ2	<u>MPSLLSQLTFWLLFTVVVEHTFASSA</u> AGTETLVGIVG...
sβ3	<u>MDTKRQRARSRSRCDVSCFRRWVTL</u> SLLLGTTPLVDSQQQDPLTMN...
sβ6 (SA)	MFKRRSTLPPRVFTGNRKTVMEREEQQRKRAPGGFLRQGQLDLLLSLMVALTAGSTNAVPAETISQSSP...
sβ7	<u>MHSGRSYSLRYPLLGVLLTVSVCVTVA</u> EIHSSDAARFHETLLSRTMELRSATSFSLDPSRQKSRNRRFLIREDERDGLGCPDEQGQEDW NNFDLLLQHNDCLDIATLGLNSRTRRRRARTRHSLAK...
sTlp1	<u>MIPKVTAMASYTLLFLLAHPVGA</u> ASSLIKEVPQELKG...
sTlp2	<u>MTKRLAVAFPLIYVLVHLCLSM</u> GHGDLRKSQNSEHRGL/INGR...
sTlp3	<u>MLIWLRSSTATVIAVLLVYRAAA</u> VASALWSVSEYRLVDSSAR...
sTlp4	<u>MFLALWLFYFTIAVSAEGA</u> ANRVRRNRVNGAVVANSDEASSVRRHRSDIILDATTVHTSNFATLDKIPRKTGD...
sTlp5	<u>MKFLPTALALWLFQEVLA</u> NDNISQATEDTTTSRHLRKPCKTKSIQATSVLTVNEN...
sTlp6	<u>MRLTMNIPQSLLSLLVWAVLLFHSSIA</u> QRLRAPHSTNDEPVAVEALPSKKL...
sDtc	<u>MRTIPFQFWTRMLLATALCQHPSVSSLVHAA</u> ADDTGTVGLSAGKL...
sDpc	<u>MFERPMCRAALLRSGQFRRGVVLROQSRRYSTSWTLLLVTA</u> VLCSTPSAYAANRRSSSQQTRQPRKATDPDDYY...

sDnaK	<u>MLRFVSRRTVLTSSCRLVETHASSCEPVSSNLLNVHNSNDRAQSGIIRSFHASAKHEILPLLAVGTIALVGRYSWKALRRMDEDWEEYEY</u> LMEQHEKQKARENPEHASRTMA...
sDrp	<u>MTMRPRTKSLYLPLYATVFAATTLIVHCAK</u> PRGKTPQRSSSNRASSAYKPSRPSRRTKTASTSEWDDILPANEELPFAFSDFDSDY FEDENDYQQRGFRSSPAGSSSLNGSSNSNDFEMPSFQ...
sPdr1	<u>MRQHSTAMSFTGLLVFGALG</u> AILMNLISIHQHLEKDEGSRRREQASTSLRKAFARNYFTPVKSIIEVPPSADKIRGSLSSAEELGEADRFAQ EPSKRTLLEEVPLLREIDFAKQSKTLSTASNRETQRNDTTSQSFAITKLLSYGNQTPKQRSDVPVEFKV...
sSec14	<u>MRFLAPRSDVGPGRISRRLLLSVTGLQSAV</u> VPVTNASVQSIDLRAEHLLFWSGLDCAVDAALHNMGPLDDLLRGGDQPTETTSE TTCVSRPMGSSPVYNN...
sαCa1	<u>MRLIAISLCCLMPCTVRCS</u> WRNIEPLHGWENDTSGTIWRMEFNPLFTSAPTSMPPTATPSDIPSSRPSFSPASPSASPSVAPSPSPS T...
sαCa2	<u>MKLIVSALFFLSIAEA</u> DLFNYGNTDTTVDGEKSYGMPNWNRECSNEDTCRGWPKFPFLVGWDAGRNTCQWCEGGPENCGFH RQSPIDLKRDRGVIGGSNEKSCPDWHWMAYKDDTCAWKDMVDEFSIERHGLRLSVPIESDGEISCVENRNGQDVRRFPRLDYSGKGF DWW...
sORF139	<u>MTGLRCVFQTFMVLALAEA</u> VRIPLLSKTKYNPLFF...
sORF261	<u>MKLSFRGCLLSALLGDAAA</u> SISSSDARLP...
sORF532a	<u>MLHCNWQLVLLAVLYVGTKA</u> SSTPGSTLMDS...
sORF534	<u>MIRMNVSKAIIIMYGQLFAVALLVTS</u> SGVSAGSTYRSTLGAWCIPRGGSDDYDT...
sPrp	<u>MKFSAFVVFFLPLAFA</u> TERSANNLPERYLNSPTGPSL...
sP4h	<u>MCQHTLTIRAQKVPPSICFPWRTVSQTMHQSPISLRSTVVCLLMITGSLQ</u> QCCLPCVQASSAANTNTGRSTRQSTPSRPPFLAHVTD NTLDRARRVALQRHEKAVAQALENYPTNQPPDSLNADATTATATGLDSYVI...
sPel	<u>MRTQGSKSHSAFVFLCAVCGLVQA</u> ASAVGADFFGLSCGNGNSGCARGRCNHDAICTPCLQSSCETAFACSLDGKYTICQHSEPHKY FLRGGVETVNRGSSLLAGICTKNACE...
sSmc	<u>MKGLMLALALWFTS</u> NDAVEGQNRFAKLFKMNAEERRNTPGAARS...

Unterstrichen, Signalpeptid nach NN; fett, SP nach HMM; normal, putatives TPL (*transit peptide-like sequence*) nach BlastP (N-terminale Extension, unkonserviert); kursiv, matures Protein; Proteinname unterstrichen, nicht durch *in vivo* Lokalisationsstudien verifiziert.

Eine Analyse putativer TPL-Sequenzen der PPC-lokaliserten Proteine aus *P. tricornutum* erfolgte ebenfalls mit TargetP und ChloroP. Dazu wurden die Primärsequenzen der Proteine beginnend mit der ersten AS nach der ermittelten Signalpeptidschnittstelle (SignalP, siehe Tab. 8.2) durch beide Algorithmen untersucht. Abbildung 8.1 und 8.2 verdeutlichen die Defizite beider Vorhersageprogramme hinsichtlich der Detektion von TPL-Sequenzen von sekundär evolvierten Organismen mit komplexen Plastiden.

```

### targetp v1.1 prediction results #####
Number of query sequences: 30
Cleavage site predictions included.
Using PLANT networks.

```

Name	Len	cTP	mTP	SP	other	Loc	RC	TPlen
sCdc48-2	905	0.042	0.810	0.010	0.312	M	3	114
sPub	737	0.042	0.902	0.024	0.073	M	1	85
sAlpha3-1	325	0.146	0.542	0.008	0.489	M	5	85
sAlpha7-1	264	0.022	0.709	0.009	0.574	M	5	61
sAlpha7-2	273	0.855	0.054	0.038	0.112	C	2	58
sBeta3	234	0.041	0.295	0.021	0.814	—	3	—
sBeta6	238	0.348	0.083	0.078	0.568	—	4	—
sBeta7	427	0.532	0.101	0.045	0.482	C	5	22
sTlp1	702	0.270	0.076	0.084	0.738	—	3	—
sTlp2	344	0.081	0.298	0.106	0.712	—	3	—
sTlp3	533	0.239	0.244	0.061	0.394	—	5	—
sTlp4	564	0.124	0.612	0.043	0.153	M	3	25
sTlp6	564	0.369	0.115	0.013	0.591	—	4	—
sDtc	470	0.045	0.134	0.266	0.859	—	3	—
sDpc	438	0.070	0.628	0.019	0.539	M	5	5
sDnaK	492	0.793	0.087	0.021	0.492	C	4	26
sPdr1	508	0.232	0.173	0.023	0.648	—	3	—
sDrp	941	0.829	0.582	0.020	0.049	C	4	26
sSec14	420	0.038	0.147	0.331	0.299	S	5	30
sAlphaCa-1	469	0.582	0.087	0.007	0.460	C	5	100
sAlphaCa-2	684	0.135	0.077	0.107	0.938	—	1	—
sORF139	171	0.040	0.139	0.160	0.523	—	4	—
sORF261	479	0.733	0.024	0.317	0.178	C	3	40
sORF532a	517	0.828	0.025	0.014	0.306	C	3	28
sORF534	290	0.025	0.365	0.045	0.865	—	3	—
sPrp	364	0.519	0.118	0.002	0.335	C	5	4
sP4h	562	0.950	0.235	0.009	0.013	C	2	55
sPel	1166	0.530	0.115	0.113	0.376	C	5	38
sSmc	653	0.043	0.413	0.030	0.656	—	4	—
TPT2	408	0.791	0.152	0.015	0.120	C	2	77
cutoff		0.000	0.000	0.000	0.000			

Abb. 8.1 TargetP-Analyse der putativen TPL-Sequenzen *in vivo* lokalisierter PPC-Proteine aus *P. tricornutum*. Die Analyse putativer TPL-Sequenzen für 30 PPC-/PPM-lokalisierte Proteine erfolgte beginnend mit der ersten AS nach der potentiellen Signalpeptidschnittstelle (SignalP, siehe Tab. 8.2) über *PLANT networks*. TargetP erkannte für 12 Proteine keine N-terminale Zielsteuerungssequenz, während für 11 Sequenzen ein mögliches plastidäres Transitpeptid, für sechs eine mitochondrielles *targeting peptide* und für ein Protein ein Signalpeptid vorhergesagt wurde. Loc: S, Signalpeptid; C, plastidäres Transitpeptid; M, mitochondrielles *targeting peptide*; -, keine N-terminale Zielsteuerungssequenz detektiert.

```

### chlorop v1.1 prediction results #####
Number of query sequences: 30

```

Name	Length	Score	cTP	CS- score	cTP- length
sCdc48-2	905	0.465	-	-0.038	36
sPub	737	0.459	-	2.361	44
sAlpha3-1	325	0.506	Y	2.014	40
sAlpha7-1	264	0.451	-	3.558	8
sAlpha7-2	273	0.532	Y	3.146	58
sBeta3	234	0.521	Y	1.183	42
sBeta6	238	0.501	Y	2.917	33
sBeta7	427	0.460	-	2.805	22
sTlp1	702	0.432	-	2.539	25
sTlp2	344	0.436	-	-0.013	86
sTlp3	533	0.440	-	1.857	15
sTlp4	564	0.453	-	6.398	14
sTlp6	564	0.477	-	-1.308	16
sDtc	470	0.435	-	5.230	16
sDpc	438	0.444	-	1.213	14
sDnaK	492	0.449	-	6.479	26
sPdr1	508	0.441	-	0.310	25
sDrp	941	0.563	Y	3.550	26
sSec14	420	0.432	-	1.867	4
sAlphaCa-1	469	0.511	Y	2.362	100
sAlphaCa-2	684	0.432	-	4.079	27
sORF139	171	0.436	-	0.393	26
sORF261	479	0.475	-	0.624	40
sORF532a	517	0.506	Y	0.765	28
sORF534	290	0.431	-	-0.906	33
sPrp	364	0.548	Y	-1.046	4
sP4h	562	0.555	Y	5.228	55
sPel	1166	0.467	-	2.360	38
sSmc	653	0.443	-	4.322	14
TPT2	408	0.583	Y	3.817	77

Abb. 8.2 ChloroP-Analyse der putativen TPL-Sequenzen *in vivo* lokalisierter PPC-Proteine aus *P. tricornutum*. Die Analyse putativer TPL-Sequenzen für 30 PPC/PPM-lokalisierte Proteine erfolgte beginnend mit der ersten AS nach der potentiellen Signalpeptidschnittstelle (SignalP, siehe Tab.8.2). Durch ChloroP konnte lediglich für 10 der 30 Sequenzen ein plastidäres Transitpeptid vorhergesagt werden. Bei den restlichen 20 Sequenzen wurde ein solches nicht erkannt. Y, plastidäres Transitpeptid; - kein plastidäres Transitpeptid.

Um potentielle Gemeinsamkeiten PPC-spezifischer TPL-Sequenzen zu analysieren wurde ein Alignment der Primärsequenzen aller bisher bekannten PPC-Proteine aus *P. tricornutum* mit den ersten 60 AS nach der vorhergesagten Signalpeptidschnittstelle (SignalP 3.0) generiert. Wie Abbildung 8.3 verdeutlicht, finden sich Gemeinsamkeiten bei den Transitpeptiden der Archaeplastida und den TPL-Sequenzen von PPC-lokaliserten Proteinen aus *P. tricornutum*. So sind bei beiden gleichermaßen hydroxylierte AS (Serin und Threonin) abundant und positive Ladungen überwiegen am N-Terminus der Zielsteuerungssequenzen (Patron & Waller, 2007). Die IDs der einzelnen Proteine sind den Tabellen unter 8.1 und 8.3, sowie Stork et al. (2012) zu entnehmen.

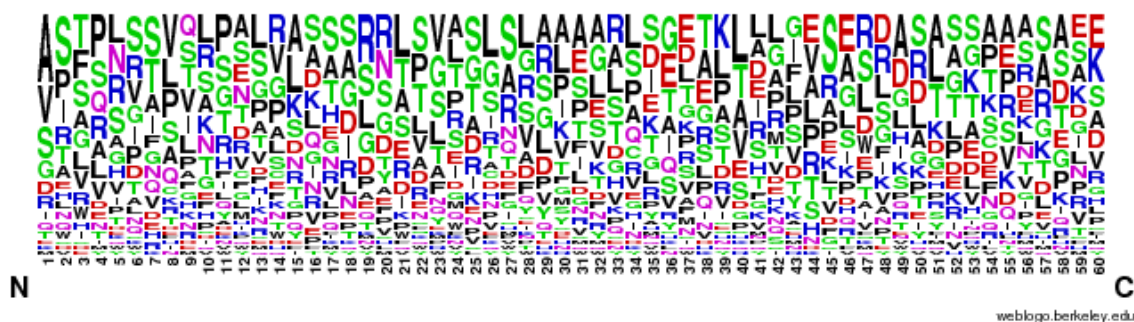
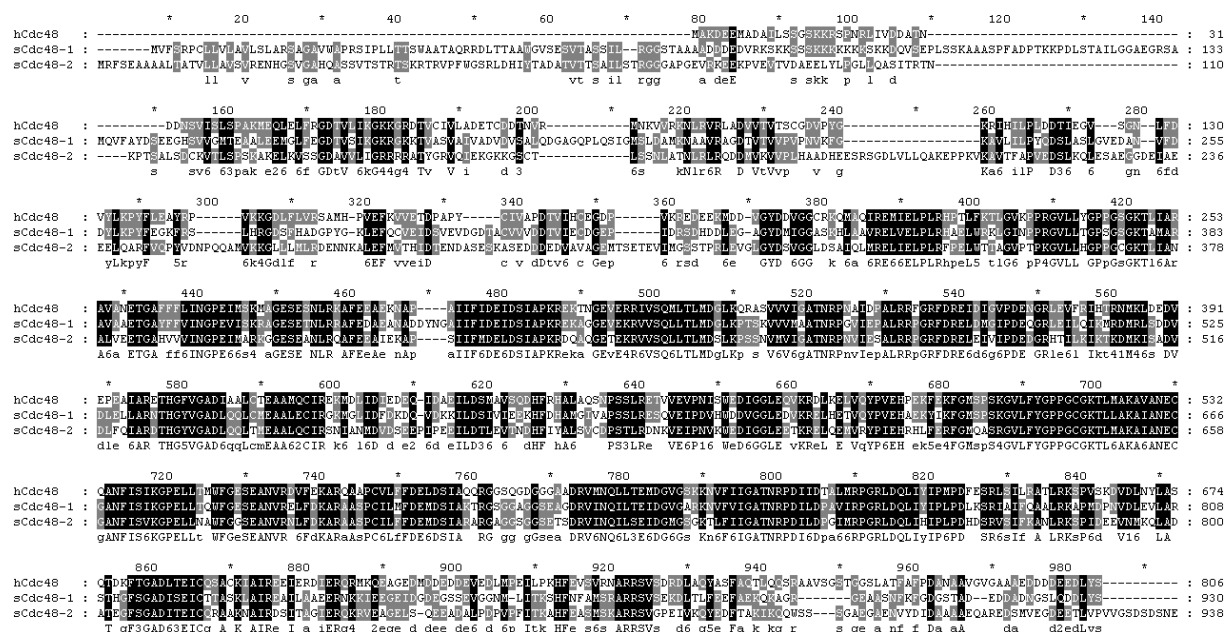


Abb. 8.3 Alignment mit relativen Aminosäure-Häufigkeiten der TPL-Sequenzen aller verifizierter PPC-Proteine in *P. tricornutum*. Das Alignment zeigt einen Vergleich der TPL-Sequenzen bekannter PPC-Proteine aus *P. tricornutum*. Position 1 entspricht der ersten Aminosäure nach der potentiellen SP-Schnittstelle (SignalP). Die Größe einer jeweiligen AS korreliert mit der relativen Häufigkeit, mit der diese an dieser Position auftritt. Zum Vergleich wurde ein Bereich von 60 AS gewählt, auch wenn dieser bei einigen Sequenzen bereits Teile des maturen Proteins umfasst. Generell kann ein gehäuftes Vorkommen hydroxylierter AS (S, T) festgestellt werden. In den ersten 35 AS finden sich häufiger positiv geladen als negativ geladene AS. Das Alignment wurde mit WebLogo erstellt (6.1.10); n = 48.

8.3. Alignments

Alle in dieser Arbeit dargestellten Alignments wurden mit ClustalX und Genedoc erstellt (6.1.10).

8.3.1. sCdc28-1/2 und hCdc48



8.3.2 sDrp und Cm_Dnm2(DRP5B)

```

Pt_sDrp      : 1 20 40 60 80 100 120
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 136
Cm_Dnm2 (DRP5B) : 1 20 40 60 80 100 120
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 89

Pt_sDrp      : 140 160 180 200 220 240 260
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 272
Cm_Dnm2 (DRP5B) : 140 160 180 200 220 240 260
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 223

Pt_sDrp      : 280 300 320 340 360 380 400
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 408
Cm_Dnm2 (DRP5B) : 280 300 320 340 360 380 400
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 352

Pt_sDrp      : 420 440 460 480 500 520 540
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 544
Cm_Dnm2 (DRP5B) : 420 440 460 480 500 520 540
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 475

Pt_sDrp      : 560 580 600 620 640 660 680
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 672
Cm_Dnm2 (DRP5B) : 560 580 600 620 640 660 680
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 611

Pt_sDrp      : 700 720 740 760 780 800
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 801
Cm_Dnm2 (DRP5B) : 700 720 740 760 780 800
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 720

Pt_sDrp      : 820 840 860 880 900 920 940
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 903
Cm_Dnm2 (DRP5B) : 820 840 860 880 900 920 940
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 855

Pt_sDrp      : 960 980 1000 1020 1040 1060
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 962
Cm_Dnm2 (DRP5B) : 960 980 1000 1020 1040 1060
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 962

```

8.3.3 sPdr1 und Cm_PDR1

```

sPdr1      : 1 20 40 60 80 100 120
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 78
Cm_PDR1_CMR358C : 1 20 40 60 80 100 120
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 135

sPdr1      : 140 160 180 200 220 240 260
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 193
Cm_PDR1_CMR358C : 140 160 180 200 220 240 260
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 270

sPdr1      : 280 300 320 340 360 380 400
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 294
Cm_PDR1_CMR358C : 280 300 320 340 360 380 400
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 405

sPdr1      : 420 440 460 480 500 520 540
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 404
Cm_PDR1_CMR358C : 420 440 460 480 500 520 540
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 538

sPdr1      : 560 580 600 620 640 660 680
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 526
Cm_PDR1_CMR358C : 560 580 600 620 640 660 680
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 673

sPdr1      : 680
               *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 528
Cm_PDR1_CMR358C : 680
               *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 683

```

8.3.4 sORF139 und der Nucleomorph-kodierte ORF139 aus *G. theta* mit dem Homolog aus *C. merolae*

```

s139      : 1 20 40 60 80 100 120 140
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 124
Gt_2oRF139 : 1 20 40 60 80 100 120 140
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 115
CMR441C    : 1 20 40 60 80 100 120 140
               * * * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 116

s139      : 160 180 200
               * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 190
Gt_2oRF139 : 160 180 200
               * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 139
CMR441C    : 160 180 200
               * * *
               MTRPRTKSLYLVLPLVATVFAATTLIVHCAPGRGRKIPSSSSNRASSASRRSPSRKRTASTSEVDDLLPAKEELPFAFDSDYFEDENDYQCGHGDSSSPAGSSSLNGGSSNSNDFFENRPTNNMPPSK : 190

```

8.3.5 sORF261 und Homologe aus *G. theta* und *C. merolae*

[illegible]

8.3.6 sORF532a und Homologe aus *G. theta* und *C. merolae*

[illegible]

8.3.7 sORF534 und Homologe aus *G. theta* und *C. merolae*

[illegible]

8.4 Proteinexpressionstests

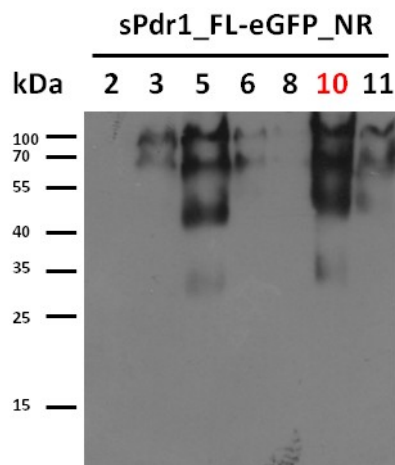


Abb. 8.4 Expressionstest von sPdr1_FL-eGFP in *P. tricornutum* unter Nitrat-induzierbarem Promotor. Gesamtproteinextrakt (6.2.3.1) von sieben unterschiedlichen Klonen, welche das PPC-lokalisierte sPdr1_FL-eGFP-Fusionsprotein in *P. tricornutum* exprimieren, wurde über SDS-PAGE und anschließenden Western Blot analysiert. Dazu wurden pro Klon (siehe Nummern) jeweils 15 µg Gesamtprotein aufgetragen und gelelektrophoretisch separiert. Die Detektion von überexprimiertem GFP-Fusionsprotein erfolgte mittels GFP-Antikörper (1:3000). Klon 10 zeigte die stärkste Expression (höchster Anteil an sPdr1_FL-eGFP am Gesamtprotein) und wurde für die weiteren Experimente verwendet.

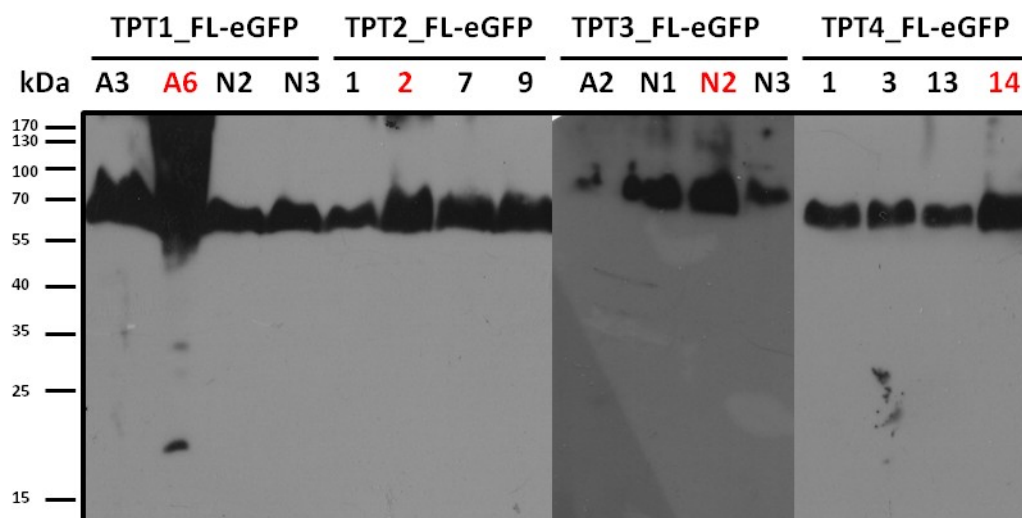


Abb. 8.5 Expressionstest von TPT1-4_FL-eGFP in *P. tricornutum* über pPha-T1. Von jeweils vier Klonen, die lichtinduziert (pPha-T1) TPT1-4_FL-eGFP Fusionsproteine in *P. tricornutum* exprimieren, wurde Gesamtproteinextrakt (6.2.3.1) hergestellt und über SDS-PAGE bzw. Western Blot analysiert. Pro Klon (siehe Nummern) wurden jeweils 25 µg Gesamtprotein aufgetragen und über SDS-PAGE separiert. Die Detektion von überexprimiertem GFP-Fusionsprotein erfolgte mittels GFP-Antikörper (1:3000). Diejenigen Klone, welche die stärkste Expression zeigten (Nummern in rot, höchster Anteil an TPT1-4_FL-eGFP am Gesamtprotein) wurden zur Carbonatextraktion (6.2.3.7) eingesetzt. Berechnete molekulare Massen (PeptideMass): TPT1_FL-eGFP: 68,5 kDa, TPT2_FL-eGFP: 73,8 kDa, TPT3_FL-eGFP: 62,5 kDa, TPT4_FL-eGFP: 68,8 kDa.

8.5 *In silico* Analysen zu putativ PPC-lokalisierten Proteinen in Organismen mit komplexer Plastide rhodophytischen Ursprungs und Rotalgen

Die folgende Tabelle beinhaltet detektierte putative Homologe zu PPC-lokalisierten Proteinen aus *P. tricornutum* in Organismen mit komplexer Plastide rhodophytischen Ursprungs im Vergleich mit potentiellen Orthologen und Homologen aus rezenten Rotalgen. Als weitere Vertreter der Heteronkontophyten wurden *Thalassiosira pseudonana*, *Ectocarpus siliculosus* und *Aureococcus anophagefferens* analysiert, sowie die Haptophyte *Emiliania huxleyi*, die Cryptophyte *Guillardia theta* und die Apicomplexen *Plasmodium falciparum* und *Toxoplasma gondii*. Vertreter der Rotalgen stellen die einzelligen Spezies *Cyanidioschyzon merolae*, *Galdieria sulphuraria*, *Porphyridium cruentum*, sowie die multizelluläre Rhodophyte *Calliarthron tuberculosum* dar. Die Genomdatenbanken (EST-Datenbank bei *P. cruentum* und partielle Genomdaten bei *C. tuberculosum*) der genannten Organismen bzw. die NCBI-Datenbank (siehe 6.1.10) wurden entweder mittels BlastP oder tBlastn nach Homologen zu PPC-Proteinen aus *P. tricornutum* durchsucht. Identifizierte Kandidaten wurden, mit Ausnahme der Sequenzen aus Rotalgen, auf die Präsenz von EST-Daten geprüft und mittels SignalP, TargetP und BlastP auf putativ PPC-spezifische Zielsteuerungssequenzen analysiert. Proteine, die bereits in früheren Studien identifiziert wurden, sind dementsprechend gekennzeichnet (siehe Legende).

Tab. 8.3 *In silico* Analysen zu putativ PPC-lokalisierten Proteinen in Organismen mit komplexer Plastide rhodophytischen Ursprungs und Rotalgen

Protein-name	He				Ha	Cr	Ap		Rh			
	<i>P. tri</i>	<i>T. pse</i>	<i>E. sil</i>	<i>A. ano</i>	<i>E. hux</i>	<i>G. the</i>	<i>P. fal</i>	<i>T. gon</i>	<i>C. mer</i>	<i>G. sul</i>	<i>P. cru</i>	<i>C. tub</i>
sDer1-1	31697 ^{2,3}	9466* ^{#10}	CBJ29494. 1 ¹⁰	64463* ¹⁰	273097* ⁶	AAK39810 .1 ²	PF3D7_14 52300 ^(*) 2,4	TGME49 094290 ^{4,8}	CMK163C 10	stig_15;Gs 21850.1 ¹⁰	Contig135 18_5 ¹⁰	IDg9967t1 10
sDer1-2	35965 ^{2,3}	28064* ¹⁰	CBJ26891. 1* ¹⁰	71775 ¹⁰	96871* ⁶	175481* ¹⁰	PF3D7_03 14300* ^{#2,4}	TGME49 017160* 4,8	CMI159C ¹ 0	stig_63;Gs 58360.1 ¹⁰	n.d.	IDg11638t 1 ¹⁰
ptE3P	48034 ⁵	n.d.	n.d.	n.d.	202611* ¹⁰	CAC27064 .1 ²	n.d.	n.d.	CMG190C 10	stig_4;Gs0 7260.1 ¹⁰	Contig519 9_1 ¹⁰	IDg13503t 1 ¹⁰
[sUba1]	54460* ²	22104* ¹⁰	CBJ32384. 1* ^{S10}	28676* ^{S10}	459406 ⁶	175495* ¹⁰	PF3D7_13 33200* ^{2,4}	n.d.	CMI034C ¹ 0	stig_39;Gs 42920.1 ¹⁰	Contig134 81_1 ¹⁰	IDg16397t 1 ¹⁰

sUbc4	34112* ¹¹	36045 ^{#10}	CBJ26141. 1 ¹⁰	n.d.	452424* ^{#1} 0	159513 ¹⁰	PF3D7_13 45500* ⁴	TGME49 095990 ⁴	CMD152C 10	n.d.	Contig843 _1 ¹⁰	IDg6863t1 10
						CAC27017 .1 ¹⁰						
						CAC27113 .1 ¹⁰						
						97471 ¹						
[sUbc6]	56431* ²	24185* ^{#10}	CBJ31773. 1 ^{S10}	22552 ¹⁰	n.d.	n.d.	n.d.	n.d.	CMS321C ¹ 0	stig_9;Gs1 5540.1 ¹⁰	Contig231 8_1 ¹⁰	IDg12422t 10
sUb	54323 ²	1539* ^{#10}	n.d.	n.d.	428400* ¹⁰	155024* ¹⁰	PF3D7_08 15700* ⁴	n.d.	CMK296C 10	stig_5;Gs1 0140.1 ¹⁰	GCDJ7DB 01DUFYC_ 5 ¹⁰	n.d.
sCdc48-1	20914 ²	40725 ¹⁰	CBN7984 7.1* ¹⁰	20739* ¹⁰	421356 ⁶	AAK39773 .1 ²	PF3D7_07 11000* ^{2,4}	TGME49 121640 ^{4,8}	CML023C ¹ 0	stig_37;Gs 41710.1 ¹⁰	Contig232 _1 ¹⁰	n.d.
sCdc48-2	19162 ⁷	4534* ¹⁰	CBJ48913. 1* ¹⁰	52190 ¹⁰	461250* ^{#1} 0	AAK39859 .1 ¹⁰	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
[sUfd1]	49319 ²	22140* ¹⁰	CBJ25981. 1* ¹⁰	30974* ^{#10}	362690 ⁶	AAF24006 .1 ²	PF3D7_09 16500 ^{#2,4}	TGME49 085700 ^{4,8}	CMK008C 10	stig_39;Gs 43460.1 ¹⁰	GCDJ7DB 01CJNOL_ 5 ¹⁰	n.d.
sPub	37661 ⁷	270310*	Esi0043_0 092*	36209 ⁷	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
ptDUP	56710 ⁵	21460 ^{S10}	CBJ48811. 1* ¹⁰	71110* ¹⁰	201710* ¹⁰	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
sHsp70	55890 ¹	263422* ^{#1} 0	n.d.	52275 ¹⁰	470246 ¹⁰	AAK39876 .1 ¹⁰	n.d.	n.d.	CMP145C 10	stig_10;Gs 16120.1 ¹⁰	n.d.	IDg8866t1 10
sDpc	48633 ⁷	n.d.	CBJ26316. 1 ¹⁰	n.d.	48584* ^{#10}	148665* ¹⁰	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
sDtc	34512 ⁷	269547* [#]	Esi0047_0 067*	18078* ^{#S}	454697	72696* [#]	n.d.	n.d.	CMR299C	stig_36;Gs 41290.1	n.d.	n.d.
sDnaK	50480 ^(*)	269683 ^U	Esi0198_0 024* ^{M?}	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
[sAlpha1]	n.d.	n.d.	n.d.	n.d.	n.d.	XP_00171 3620.1 ¹⁰	n.d.	n.d.	CMN208C 10	stig_15;Gs 21750.1 ¹⁰	GCDJ7DB 01E2ROP_ 2 ¹⁰	n.d.
[sAlpha2]	bd_206* [#]	22280* ^{#10}	CBJ32748. 1 ¹⁰	30228* ¹⁰	n.d.	XP_00171 3542.1 ¹⁰	n.d.	n.d.	CML273C ¹ 0	stig_21;Gs 27920.1 ¹⁰	Contig261 2_3 ¹⁰	n.d.

sAlpha3 (3-1)	45475 ^{#7}	9834* ^{#10}	CBN7576 4.1* ¹⁰	n.d.	n.d.	XP_00171 3481.1 ¹⁰	n.d.	n.d.	CMR302C ¹⁰	stig_11;Gs 16900.1 ¹⁰	Contig179 5_2 ¹⁰	IDg16652t 1 ¹⁰
[sAlpha3 (3-2)]	42884	10903 ^{#10}										
[sAlpha4]	n.d.	n.d.	CBJ27439. 1* ¹⁰	n.d.	n.d.	XP_00171 3416.1 ¹⁰	n.d.	n.d.	CMT379C ¹⁰	stig_21;Gs 27820.1 ¹⁰	Contig623 1_6 ¹⁰	IDg16723t 1 ¹⁰
[sAlpha5]	n.d.	n.d.	CBN7673 4.1* ¹⁰	n.d.	n.d.	XP_00171 3546.1 ¹⁰	n.d.	n.d.	CMQ225C ¹⁰	stig_10;Gs 15880.1 ¹⁰	Contig856 8_5 ¹⁰	n.d.
[sAlpha6]	n.d.	n.d.	CBJ48831. 1 ^{#10}	22897* ^{#10}	n.d.	XP_00171 3620.1 ¹⁰	n.d.	n.d.	CME128C ¹⁰	stig_55;Gs 54410.1 ¹⁰	Contig561 5_2 ¹⁰	IDg8087t1 10
sAlpha7 (7-1)	17388 ⁷	n.d.	n.d.	n.d.	n.d.	XP_00171 3184.1 ¹⁰	n.d.	n.d.	CMO197C ¹⁰	stig_14;Gs 20880.1 ¹⁰	Contig821 1_4 ¹⁰	IDg14765t 1 ¹⁰
sAlpha7-2	43079 ^{#7}	6655 ¹⁰	n.d.	n.d.	n.d.	120764* ¹⁰	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
[sAlpha]	salpha ^{#10}											
sBeta1	38228 ¹⁰	8470* ^{#10}	n.d.	n.d.	n.d.	AAK39875 .1 ¹⁰	n.d.	n.d.	CMH154C ¹⁰	stig_32;Gs 38030.1 ¹⁰	Contig130 6_3 ¹⁰	n.d.
sBeta2	35028 ⁷	22550* ¹⁰	n.d.	n.d.	n.d.	CAC27046 .1 ¹⁰	n.d.	n.d.	CMK289C ¹⁰	stig_6;Gs1 1110.1 ¹⁰	Contig112 24_1 ¹⁰	n.d.
[sBeta3]	12731*	7983* ¹⁰	n.d.	52492 ¹⁰	n.d.	AAK39755 .1 ¹⁰	n.d.	n.d.	CMG188C ¹⁰	stig_4;Gs0 7240.1 ¹⁰	Contig171 9_4 ¹⁰	n.d.
[sBeta4]	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	CMH119C ¹⁰	stig_23;Gs 30300.1 ¹⁰	Contig805 6_1 ¹⁰	n.d.
[sBeta5]	n.d.	n.d.	CBN7549 5.1 ^{#10}	20025* ^{#10}	n.d.	AAK39885 .1 ¹⁰	n.d.	n.d.	CMO149C ¹⁰	stig_12;Gs 19040.1 ¹⁰	Contig317 8_3 ¹⁰	IDg10070t 1 ¹⁰
sBeta6	49432 ⁷	3566 ¹⁰	CBN8022 0.1* ¹⁰	n.d.	n.d.	CAC27058 .1 ¹⁰	n.d.	n.d.	CMT477C ¹⁰	stig_70;Gs 60610.1 ¹⁰	Contig135 56_6 ¹⁰	n.d.
sBeta7	22110 ⁷	n.d.	n.d.	n.d.	n.d.	AAK39749 .1 ¹⁰	n.d.	n.d.	CMD060C ¹⁰	stig_31;Gs 37470.1 ¹⁰	Contig157 64_2 ¹⁰	IDg12078t 1 ¹⁰
sTlp1	23414 ⁷	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
sTlp2	35803 ^M	11395*	n.d.	55219	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
sTlp3	40345 ^M	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
sTlp4	40553*	n.d.	n.d.	n.d.	n.d.	85246* [#]	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
[sTlp5]	44903	256513 [#]	n.d.	n.d.	194015 ^U	n.d.	n.d.	n.d.	n.d.	n.d.	Contig281 6_3	IDg10886t 1
		40831			193990						Contig142 91_2	IDg20785t 1M

		266343 ^U			226937*							IDg18826t 1
		23662			238496* ^U _M							
		263452 [#]										
		42804										
sTlp6	49698	9289*	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.		IDg11875t 1
sSec14	44766 ⁷	1526* [#]	Esi0000_0 552* ^{M?}	n.d.	452613	n.d.	n.d.	n.d.	CMT608C *	stig_38;Gs 42470.1*	Contig964 5_6	n.d.
sDrp	373797 ^(*)	6025*	Esi0121_0 052*	60924*	96578*	n.d.	n.d.	n.d.	CMN262C	n.d.	n.d.	n.d.
sPdr1	40314 ^(*)	n.d.	Esi0010_0 059* ^{M?}	scaffold_4 73744- 474964*	445242 ^M	141876*	n.d.	n.d.	CMR358C *	n.d.	n.d.	IDg18795t 1 [?]
				scaffold_5 :491946- 493125*								
sORF532a	45935 ⁷	23216*	Esi0090_0 039 [#]	scaffold_2 9:368492- 450807*	268225*	XP_00171 3555	PF3D7_06 24400 [#]	n.d.	CMQ054C	stig_42;Gs 45770.1*	Contig352 5_5	IDg11803t 1
sORF534	48879 ^{(*)7}	269759*	CBJ32213 9	21663* [#]	464972	XP_00171 3632	PF3D7_12 12100	TGME49 087270* ^{S9}	CMI197C	stig_59;Gs 56340.2*	GCDJ7DB 01A44EB_ 3	n.d.
sORF261	47811 ⁷	24247	Esi0234_0 011	n.d.	114318* ^U	XP_00171 3372	n.d.	n.d.	CMD101C	stig_27;Gs 33600.1*	Contig118 81_3	IDg15730t 1
												IDg17166t 1
sORF139	31704* ⁷	2280*	n.d.	n.d.	n.d.	XP_00171 3268	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
sSmc	42675 ⁷	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	CMQ429C ?	stig_4;Gs0 7740.1* [?]	n.d.	n.d.
sPel	37424* ⁷	23888* [#]	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	CMI068C	stig_421; Gs66180. 1*	Contig136 27_3 [?]	n.d.

sP4h	47766 ^{*7}	21399	n.d.	n.d.	467430	147722 ^{*U} _M	n.d.	n.d.	n.d.	n.d.	Contig491 1_3 [?]	n.d.
sPrp	41316 ⁷	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
sαCa1	35370 ^{*7}	22391	Esi0000_0 144	63566 ^{*#M?}	n.d.	135711 ^{*M}	n.d.	n.d.	n.d.	n.d.	n.d.	IDg12365t 1
						138994 [*]						
sαCa2	42574 ⁷	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
s6PGDH	45333 ¹²	n.d.	n.d.	65719 ^{*#}	n.d.	n.d.	n.d.	n.d.	CMS195C	stig_1;Gs0 2840.1 [*]	Contig117 86_2	IDg20067t 1
sTrxH	48539 ¹³	5491 [*]	Esi0003_0 293	17366 ^{*#}	74645	153526 ^{*U}	PF3D7_09 16100 [?]	TGME49 070120 [?]	CMJ096C	stig_8;Gs1 3580.1 [*]	Contig122 41_3	IDg12852t 1
					471051	84862 ^{*U}						
sNTRC	37815 ¹³	40200 ^U	Esi0030_0 003 ^{S?}	13221 [#]	64001 ^{#S}	72625 [*]	n.d.	n.d.	CMI298C	stig_32;Gs 37990.1 [*]	Contig468 2_3	n.d.

He, Heterokontophyta; Ha, Haptophyta, Cr, Cryptophyta; Ap, Apicomplexa; Rh, Rhodophyta; *P. tri*, *Phaeodactylum tricornutum*; *T. pse*, *Thalassiosira pseudonana*; *E. sil*, *Ectocarpus siliculosus*; *A. ano*, *Aureococcus anophagefferens*; *E. hux*, *Emiliania huxleyi*; *G. the*, *Guillardia theta*; *P. fal*, *Plasmodium falciparum*; *T. gon*, *Toxoplasma gondii*; *C. mer*, *Cyanidioschyzon merolae*; *P. cru*, *Porphyridium cruentum*; *C. tub*, *Calliarthron tuberosum*.

, 5' keine EST-Abdeckung; () Genmodell experimentell verifiziert (cDNA); #, Genmodell unterscheidet sich von der kodierten Proteinsequenz, welche für die Analyse N-terminaler Zielsteuerungssequenzen verwendet wurde (SignalP, TargetP, BlastP); U, Startcodon/EST-Datenlage unklar; S, kein identifizierbares Signalpeptid (SignalP 3.0), Genmodell möglicherweise unvollständig, Protein aber aufgrund von Homologie zu *P. tricornutum* PPC-Protein höchstwahrscheinlich symbiontspezifisch; M, potentiell Membranprotein (TMHMM); ?, Orthologie bzw. Homologie unklar; n.d., nicht detektiert, homologes Protein entweder nicht vorhanden oder aufgrund evolutionärer Distanz mit der Sequenz des PPC-Proteins aus *P. tricornutum* nicht detektierbar; Proteinname, PPC-Protein aus *P. tricornutum*; [Proteinname], Lokalisation des *P. tricornutum* Proteins nicht über *in vivo* Lokalisationsstudien verifiziert; die IDs verweisen auf die zugehörigen Proteinsequenzen in den Genomdatenbanken der untersuchten Organismen bzw. der NCBI-Proteindatenbank (6.1.10). Quellen gekennzeichnete Protein-IDs:

¹: Gould SB, Sommer MS, Kroth PG, Gile GH, Keeling PJ, Maier UG (2006) *Mol Biol Evol* 23:2413-2422. ²: Sommer MS, Gould SB, Lehmann P, Gruber A, Przyborski JM, Maier UG (2007) *Mol Biol Evol* 24:918-928. ³: Hempel F, Bullmann L, Lau J, Zauner S, Maier UG (2009) *Mol Biol Evol* 26:1781-1790. ⁴: Spork S, Hiss JA, Mandel K, Sommer M, Kooij TWA, Chu T, Schneider G, Maier UG, Przyborski JM (2009) *Eukaryot Cell* 8:1134-1145. ⁵: Hempel F, Felsner G, Maier UG (2010) *Mol Microbiol* 76:793-801. ⁶: Felsner G, Sommer MS, Gruenheit N, Hempel F, Moog D, Zauner S, Martin W, Maier UG (2011) *Genome Biol Evol* 3:140-150. ⁷: Moog D, Stork S, Zauner S, Maier UG (2011) *Genome Biol Evol* 3:375-382. ⁸: Agrawal S, van Dooren GG, Beatty WL, Striepen B (2009) *J Biol Chem* 284:33683-33691. ⁹: Sheiner L, Demerly JL, Poulsen N, Beatty WL, Lucas O, Behnke MS, White MW, Striepen B (2011) *PLoS Path* 7:e1002392. ¹⁰: Stork S, Moog D, Przyborski JM, Wilhelmi I, Zauner S, Maier UG (2012) *Eukaryot Cell* doi:10.1128/EC.00183-12. ¹¹: Bullmann L, Haarmann R, Mirus O, Bredemeier R, Hempel F, Maier UG, Schleiff E (2010) *J Biol Chem* 285(9):6848-56. ¹²: Gruber A, Weber T, Bartulos CR, Vugrinec S, Kroth PG (2009) *J Basic Microbiol* 49(1):58-72. ¹³: Weber T, Gruber A, Kroth PG (2009) *Mol Plant* 2(3):468-477.

Danksagung

Mein besonderer Dank gilt Prof. Dr. Uwe Maier, der mir die Möglichkeit gegeben hat an einem überaus interessanten und abwechslungsreichen Thema arbeiten zu können, sowie für sein stetiges Interesse am Fortgang meiner Arbeit und die vielen ertragreichen Diskussionen und Ideen. Außerdem möchte ich mich bei ihm für die Bereitstellung des Arbeitsplatzes und der finanziellen Mittel bedanken.

Des Weiteren gilt mein Dank Prof. Dr. Andreas Brune für die Erstellung des Zweitgutachtens zu dieser Arbeit, sowie Prof. Dr. Renate Renkawitz-Pohl und Prof. Dr. Alfred Batschauer als weiteren Mitgliedern der Prüfungskommission.

Vielen Dank an Dr. Stefan Zauner und Dr. Christopher Grosche für das Korrekturlesen meiner Arbeit. Stefan sei darüberhinaus für viele konstruktive Diskussionen und Ratschläge im Laboralltag gedankt.

Ich danke der International Max Planck Research School (IMPRS) und dem Max Planck Institut für terrestrische Mikrobiologie in Marburg für die Finanzierung und Möglichkeit zur Teilnahme an Seminaren, Workshops und internationalen Konferenzen.

Ein ganz besonderes Dankeschön geht an Simone Stork und Ilka Wilhelmi für die Unterstützung beim PPC-Projekt im Rahmen ihrer Masterarbeiten, insbesondere in den Bereichen der Plastidenteilungsfaktoren, proteasomalen Komponenten und Proteasen.

Vielen Dank an Dr. Franziska Hempel und Dr. Lars Abram für die Zurverfügungstellung von Self-Assembling- und Split-GFP-Konstrukten.

Ich danke allen aktuellen und ehemaligen Mitgliedern der AG Maier für das herausragende Arbeitsklima und die schöne Zeit auch abseits der Wissenschaft.

Letztendlich möchte ich noch ganz besonders meiner Familie für ihre Unterstützung danken.

Curriculum Vitae

Persönliche Daten	Daniel Moog geboren am 06.04.1984 in Marburg
Promotion 10/2009-12/2012	Philipps-Universität Marburg, Fachbereich Biologie, Promotion im Fachgebiet Zellbiologie, Arbeitsgruppe: Prof. Dr. Uwe Maier Titel der Dissertation: „Charakterisierung eines minimierten eukaryoten Cytoplasmas – das periplastidäre Kompartiment der Diatomee <i>Phaeodactylum tricornutum</i> “
Hochschulausbildung 09/2007-09/2009	Philipps-Universität Marburg, Fachbereich Biologie, Masterstudium mit den Schwerpunkten Zellbiologie, molekulare Mikrobiologie, mikrobielle Ökologie und Biochemie, Masterarbeit im Fachgebiet Zellbiologie, Arbeitsgruppe: Prof. Dr. Uwe Maier Titel der Masterarbeit: „Heterologe Lokalisationsstudien symbiontspezifischer ERAD-Faktoren aus <i>Emiliana huxleyi</i> in <i>Phaeodactylum tricornutum</i> und PPC-Proteomik in <i>Phaeodactylum tricornutum</i> : Analyse eines putativ Vesikel-vermittelten Proteintransports und Identifizierung PPC-lokalisierter Faktoren“
09/2004-09/2007	Philipps-Universität Marburg, Fachbereich Biologie, Bachelorstudium, Bachelorarbeit im Fachgebiet Zellbiologie, Arbeitsgruppe: Prof. Dr. Uwe Maier Titel der Bachelorarbeit: „Identifizierung und Charakterisierung von ERAD-Komponenten in <i>Emiliana huxleyi</i> “
Grundwehrdienst 10/2003-06/2004	6. Artillerieaufklärungsbataillon 131 in Neustadt (Hessen)
Schulische Ausbildung 1994-2003	Besuch des Gymnasiums Lahntalschule in Biedenkopf

Erklärung

Hiermit versichere ich, dass ich meine Dissertation

„Charakterisierung eines minimierten eukaryoten Cytoplasmas - das periplastidäre Kompartiment der Diatomee *Phaeodactylum tricornutum*“

selbstständig, ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 03.12.2012

Daniel Moog